Listening To Music

The Development of a Technique to Evaluate the Quality of Responses to Music Using the SOLO Taxonomy

oriel

Helen O.Lee B.Ed (Hons)

Submitted in fulfilment of the requirements
for the degree of Master of Education
University of Tasmania
1991

ABSTRACT

Music Education throughout the world recognises three main areas of activity : Listening, Performance and Composition, and requires that they all be evaluated effectively. This study, which began as an attempt to devise an effective evaluation tool for moderating standards in the Listening area between schools, has developed an evaluation technique for assessing written responses to music which can be used for any type of music in any classroom situation. It does not separate music listening experiences into separate elements but deals with the total gestalt of the listening to music experience.

The test used in developing the technique consisted of three extracts from music written for orchestral instruments which were played three times to students (Grades 7-10, aged 12-16) An open ended question was set to which students gave a written response in their own words. Responses were analysed for the musical elements mentioned and levels of thought were revealed using the SOLO Taxonomy.

The SOLO Taxonomy (Biggs and Collis 1982) provides a mechanism to evaluate the quality of learning by examination of the Structure of the Observed Learning Outcomes (SOLO). Student responses to classroom tasks are classified into five levels of increasing complexity of thought : Prestructural, Unistructural, Multistructural, Relational and Extended Abstract.

The most complex level, Extended Abstract, is a recognition of new, flexible, original thinking. Each of the levels reveals a new stage in comprehension of the implications of the task and of thinking about it, and they can be applied to almost any topic. In order to stimulate the higher levels, Relational and Extended Abstract, problems must be devised which do not have instant one word solutions, but which require the activation of deeper thought processes through recall of previous knowledge, comparison, discrimination, recognition, clarification, classification, review and restructuring of knowledge.

Trials of the test materials were made in a one year Pilot Study, and 328 students in two Tasmanian High Schools tested in the three year Main Study produced 1260 individual responses. Some students were tested twice, and after SOLO-based tuition, a small group was tested for a third time. The influence of Written Fluency, Music Listening Ability, Performance ability and Motivation upon responses was examined. Comparisons were made with response assessments by nine experienced Music teachers, who also classified responses into SOLO levels.

ACKNOWLEDGEMENTS

I should like to record my sincere appreciation to all those who so generously gave me their time and experience during the preparation of this thesis.

To Professor Kevin Collis I express my gratitude for his willingness to allow me to reshape SOLO for my needs, for his invaluable advice and expert guidance, and his patient support throughout this project.

I would also like to record the help of other staff and students at the University of Tasmania, and in particular to the staff of all the university libraries which I used. I was particularly impressed by the professional attitude, interest, and willingness to help a stranger shown by librarians in Auckland, London and Reading.

The comradeship and advice given me by music researchers at ISME and ASME conferences was also a major source of inspiration and encouragement to continue this study whilst teaching full time.

This study would not have been possible without the cooperation of my colleagues and students. My colleagues generously gave their time to discuss the project at a time of great industrial turmoil in schools, encouraged me to continue with the project, and made its administration possible. The students, who trialled so many different procedures and listened so carefully for me, also provided inspiration as I observed their growth as listeners.

Finally I must thank my husband, Maurice, for his inspiration, his patience, and his unfailing support throughout the project.

DECLARATION

I herewith declare that this thesis contains no material which has been accepted for the award of any other degree or diploma in any university, and that, except where due reference is made in the text, to the best of my knowledge and belief it contains no material previously published or written by another person.

H.O.Lee

Contents

Chapter Four Pilot Study

Chapter Five Main Study

Chapter Six Analysis of the Data from the Main Study

Chapter Seven
An Examination of the SOLO Music Listening Evaluation

Chapter Eight Conclusions and Implications

Tables

Figures

Chapter One

Introduction

This chapter explains why it was considered necessary in this study to focus upon the Listening aspect of the Tasmanian Music course. Decisions made about the direction of the study were based upon consideration of the following factors : the role of music in society, the three major aspects of music, the school music curriculum with its attendant evaluation needs, and the current music syllabus evaluation practices in Tasmania.

The Role of Music in Society

Music is amongst the earliest experiences of the human species. Musical stimuli are amongst the first responded to by babies, and music is used naturally by young children as a medium of expression for communication with each other, as signals during games and for private musings in their own songs (Wilson 1988, Gardner 1981). Music is perceived as a total gestalt by the right side of the brain, yet the creation of music however requires analysis and evaluation by the left side, so Music can be said to be a total brain activity and a suitable medium for the education of the entire brain (Woods 1990). Music can also stimulate innovative thinking, communicate values, feelings and ideas, create a sense of cultural identity and bring groups together in cooperative endeavours; and for these reasons it has always been considered to be an essential component of man's life.

Music is part of the life of all human cultures. There is even anthropological evidence that the mouth and tongue muscles essential for spoken language developed many thousands of years later than those controlling the lungs and vocal cords, thus suggesting that vocal music may have preceeded speech (Menuhin and Davis 1979). Music is used by societies to stimulate collective emotions and collaborative actions, to draw a society together and to define groupings within that society; and each type of usage requires a different style of music. Sacred and secular ceremonies and processions need solemn and uplifting music; battles require loud repetitive group sounds and frightening noises and cries; everyday tasks need rhythmic work songs and calming lullabies; long distance communication is served by slow echoing alpenhorn tunes, intricate African drumming and epic ballads. Many societies have used music, songs and dance as teaching media to convey knowledge of survival skills from one generation to the next, and this usage has recently been revived by Australian Aborigines using modern music to teach skills essential for survival in modern Australia (CAAMA 1987).

Music as an Aural Experience

Whatever the intention behind the music, it itself remains the same: the art of organising sounds into pleasing patterns. Musical traditions may vary around the world but they maintain the same basic principles; the organisation of sounds with pitches, durations, timbres and dynamics into accepted formal structures, which are then performed and participated in by both active and passive listening audiences (Jorgensen 1987). Music is essentially an aural experience, deeply thought about, planned and executed with care by both composers and performers, yet it is unable to be fully experienced without being heard. Skilled musicians claim to be able to realise the sound internally from written notation, yet they cannot communicate their music to other people without direct aural transmission. Music is an art form that exists only in sound and time. It can be partially mapped on paper, but unlike visual art, the visual images of music are not totally accurate and complete

images of the creator's intentions, but are merely indications of the directions the music should take in performance. No matter how detailed the score each performance will differ in some degree according to the understanding of the director and performers - unless of course the performance is by pre-programmed computer or is a recording. It could be argued that if Music was not performed, heard and listened to, it would not exist. Listening is a vital aspect of the musical experience.

Listening has been accepted as one of the three essential components of Music and a complete Music Education curriculum (Paynter 1982, Swanwick 1982). Every piece of music has separate elements and its uniqueness depends on the relationship between them, but the essence of the art form called Music is that there is an intrinsic interrelationship between the act of devising the sounds and their progression (Composition), the act of recreating those sounds vocally or with the aid of technological devices (vocal or instrumental Performance) and the act of acknowledging the existence of the art form by appreciating it (Listening).

The act of Composing involves close intensive listening and some performance. It is an activity stimulated by the right brain, which is then refined through recall and analysis of previous comparable passages, evaluated and rearranged into new patterns of sounds. It is important in the development of self esteem, as it places the student in the roles of both creator and judge, and it also requires the ability to step outside oneself and be objective about the products created. Composition is a highly complex cognitive and affective activity, inseparable from listening and performance.

Performance is an also an act of creation, though this time one of re-creation, through which the composer's musical intentions become the performer's personal property. The performer develops a sense of personal involvement and ownership, power over the aural environment, a sense of self esteem and worth, and through performance to others self confidence. Listening is also an essential component of performance, as not only are the accuracy of

pitch, rhythm and intonation continually monitored aurally, but decisions about the details of interpretation, such as articulation and alterations in dynamics, are also determined aurally.

Listening is an intensive skill, through which awareness of musical elements of varying cognitive levels are learned, developed and practised, so that the natural subconscious perception of simultaneous aural stimuli is brought into the area of conscious perception. It is an integral component of all musical activities and the sole means through which music can be perceived. The source of all evaluations and judgments about music, Listening is also for most people the origin of the value of individual pieces of music.

Listening and its Roles

"Listening is a creative act ; a process by which we make the sounds part of ourselves and so assimilate their meaning for us. Active listening is fundamental at every stage of improvisation, composition, interpretation and presentation...It is the means by which music is remade within anyone who properly and positively hears it...it is basic to musical experience"

John Paynter (1982)

Listening has such importance because Music is an ephemeral art, it exists only as sound. It is true that, the composer may set out the notes on a printed score, but no matter how careful the directions as to tempo, dynamics or duration, each performer will interpret them according to his views of their exact meaning, and unless each direction is programmed into a digital synthesizer, every performance will vary in minute details. In 1938 Seashore proved that there was a huge variation between performances of the same music in his laboratory-studio, where he measured the exact lengths and pitches of notes played and sung, and recorded these measurements on elaborate charts. He showed clearly the difference between performances of the same music, but he could not explain scientifically, why one performance was accepted as being better than another.

Seashore's work showed that the variations between performances might be minute, but that they existed and could be recorded. Such intensive listening is not normally required as music listening is carried out in a variety of situations for different purposes : audience listening, performance listening, composition listening, dance listening, rehearsal and work-in-progress listening. It can be argued that all of these types of listening are evaluative if carried out in an active manner.

Active listening for continual evaluation and the subsequent adjustment is essential in all areas of musical activity. Without it performances would suffer from bands and choirs out of tune, incorrectly rendered and harmonised tunes, and unsynchronised dancers. In fact, Swanwick (1982) believes that it is so important that he places listening (or in his phraseology "audition") as the cornerstone of his plan for a balanced music curriculum. This curriculum proposes that the study of musical repertoire and craftsmanship should be linked with the processes of skills development, through this essential component Audition.

Listening is arguably one of the most important tools of the musical craftsman, an essential skill to be developed by all musicians, whether performers or composers. Every musician or musically educated person of every culture is always engaged in a process of making inner aural comparisons and evaluations between the current listening experience and previous ones. Throughout rehearsals and performances musicians continually appraise their progress, listening closely to the music they produce to determine the accuracy of the notes in pitch, intonation, rhythm, articulation, dynamics, phrasing and tone colour. The decisions they make about how to alter these aspects are based upon their listening experiences and lead to personal interpretations, which may maintain, or consciously deviate, from a stylistic tradition. The intensity with which a performer listens to his own performance is one of the differences between the top professional and the less critical amateur (Galway 1986). There is a continual process of experimentation and adjustments being made on the basis of aural

judgments, which may change to suit the circumstances, as when a concert soloist insists on rehearsing in an unfamiliar hall to judge the acoustics and revise performance strategies accordingly.

Composers may hear their music inside their head using their "inner ear" or may use an instrument such as a keyboard or computer to try out their music, but in each case they need to listen critically to their music so that they can make decisions about the details of melodic shape, tonality, harmony, formal structure, texture and orchestration. Listening is an essential skill for both performers and composers.

Listening is not only essential for practising musicians, but part of everyday life. Of those students who study music at school, some, once they cease having formal lessons, continue to perform either alone at home or in self run or organised groups, and a very small number continue to compose. Everyone, whether or not they have had the opportunity to study music, continues to listen for the rest of their life. Listeners are also essential for the economic survival of both performers and composers. Concert halls must be filled and paid for, records, tapes and compact discs must be bought, and if Music is to continue to develop rather than stagnate, these listeners must be intelligent and critical enough, to demand quality, to reject inferior, repetitious musical concepts, and to be willing to experiment with fresh interpretations of old favourites and new styles of music.

If listening is so very important, then it must be taught effectively. In order to be able to do this, teachers need first to know exactly what it is they are trying to teach, and then to be able to evaluate the progress of their students, so that the success of the teaching programme can be estimated and teaching strategies revised and improved.

What is it then that music teachers are trying to teach under this heading? What is listening? What are listening skills? What is a good listener? How is it possible to recognise a good listener? According to the Macquarie Dictionary : Listening is *"to give*

attention with the ear, to attend closely for the purpose of hearing" and Hearing is *"the faculty or sense by which sound is perceived, the act of perceiving sound"*. These definitions equate with the two types of listening discussed by musicians : active and passive listening. Passive listening is an activity which people carry out continuously without being aware of doing so, hearing sounds without being aware of them, as in a supermarket or at a film. Music carries many cultural messages and passive listeners are vulnerable to them all. Active listening involves awareness of sounds, knowledge of how they are created and what their intentions are. Active listeners could be said to be musically literate listeners.

Three factors create the essential difference between the passive and active listener. These are attention, discrimination and consciousness (Hopkins 1979). The active listener is attentive to the sounds, and notices changes within the piece; is aware of alterations in melody and harmony; recognises changes in instrumentation; discriminates between instruments, dynamics, articulations, methods of phrasing, and styles of music and interpretation; and is constantly engaged in comparing the present listening experience with those that have been previously heard and evaluated. The active listener is consciously making aural comparisons.

There are many features of music that can be noted by an active listener. Those common to music of all cultures are pitch, dynamics, rhythm, timbre and form (Jorgensen 1987). Other elements found in Western musics include: tempo, melodic shape, phrasing, articulation, harmony, tonality, texture, style, and orchestration. A person who is skilled at listening to music has the ability to identify the elements used to create the sounds, to understand how their interaction produces particular effects and to communicate this understanding effectively to someone else. This skill can operate at various levels, from the simple level :
"There's a loud bit in the middle because there are more brass instruments and drums playing then," to the more academic level:
"The changes in orchestration from predominantly high woodwind and

strings to deeper strings and low brass, combine with urgent upwards phrases and a more open texture at this point to develop the crescendo into a powerful climax."
The better educated the listener, then the wider and more extensive the sound memory and musical vocabulary available for the aural comparisons necessary to make informed judgements about unfamiliar pieces of music and fresh interpretations of familiar pieces.

An educated listener presupposes the existence of an effective musical education and Music Education curricula around the world have come to focus on the intrinsic interrelationship between the three major roles of people in musical life : Listener, Performer and Composer. They are also coming to require that a balance between the three should be maintained throughout music education; and to demand that all three must be fully and equably evaluated.

An essential part of the teaching process, evaluation should be used to measure the success of a teaching programme devised to implement sound educational ideas, but all too often it becomes the force which determines the content of the curriculum. Syllabus aims can be ignored and the focus of teaching turned towards the learning of a body of knowledge, which is factual and easily tested, rather than remaining centred on the devlopment of skills and understanding, which are much harder to evaluate. Unless suitable evaluation methods are available, a process based course can be distorted to serve the needs of a content based test. The type of test used should always be selected according to the aims of the curriculum and the role to be played by the results of the individual evaluations.

The Role of Evaluation in Education

Evaluation has several distinct roles in education. Firstly, it is a classroom diagnostic tool, used formatively to assess progress and achievements made in the process of learning so that improvements and greater efficiencies can be incorporated into subsequent teaching strategies. Every conscientious teacher

continuously revises lesson plans, units of work and overall programmes as the result of their use of this type of ongoing evaluation, which may be carried out formally or by daily observations.

The second role of evaluation is often of less importance to teachers, but more important in the eyes of students, parents and the community. This is the summative assessment of the learning end product, which is all too often used to compare students with each other, to rank them and often to determine their future careers. This final endpoint evaluation of achievement should therefore be as accurate and meaningful as possible. It should also be a measure of an accepted, universally understood standard.

Evaluation in schools is undertaken to fulfill these two roles and also to act as a predictor of future progress. The methods used to obtain these evaluations are often the same, but the uses to which the results are put are not. Whereas formative and predictive assessment results are usually kept within the school, summative assessment is more public and often carried out under the authority of an official examining body.

High School Music Course Evaluation in Tasmania

In Tasmania, courses for students in Grades 9-12 are regulated and evaluated under the auspices of the Schools Board of Tasmania. Equivalent standards between schools are maintained by externally assessed components in Grades 11 and 12 in addition to the meetings of subject teachers, known as Moderation meetings, which are held regularly to discuss the interpretation of the syllabus and the standards of the internally assessed components of the syllabus in Grades 10, 11 and 12. (See Appendix A for a brief outline of the Tasmanian Education System.)

Until 1990, Music courses, like other subjects, were not formally assessed and certificates awarded until the end of Grade 10, and the moderation of standards between schools first occurred

at that stage. In the Grade 10 Music course the three aspects of Music : Performance, Composition and Listening, had been regarded as equal for many years, and Music Moderation meetings had looked at methods used to evaluate each of them.

Over the years an effective procedure for moderating Performance standards had been developed through the presentation of student solo and ensemble performances recorded on audio and video tape, which represented each of the award categories and were representative of the variety of instruments studied. The compositions, arrangements and improvisations required by the Composition section of the syllabus had also been judged through a similar examination of notated scores and recorded performances.

The Listening aspect of the course, however, was, possibly due to the lack of time, often treated in a cursory fashion at these meetings. This part of the meeting usually consisted of comparing aural tests and written assignments about composers. Occasionally a few teachers produced listening tests that required students to identify instruments, forms, style, periods and composers. The required responses to tests presented were usually only one word or a short sentence and it was difficult to see how these or the assignments could be held to be assessing the stated aims of the syllabus, which were :

"to help students respond aesthetically to music"

"to appreciate music of all idioms and styles"

or how they fitted with the evaluation direction :

"Teachers will look for evidence that students can : identify, describe and compare features of music, and show an informed awareness of design and techniques used in creating an overall effect."

It seemed as if the teachers were testing what was testable, rather than striving to develop evaluation procedures for the more complex learning of the listening skills required by the syllabus.

There seemed to be a dearth of evaluative procedures revealing evidence of students having heard and listened carefully to actual pieces of music, instead the concensus seemed to be that students

should be able to identify various elements of music and know facts about music and musicians, rather than understand or discuss their music. The wording of the questions presented -*"What? Who? Where? When?"*, required answers at the basic levels of knowledge and comprehension (Bloom, Hastings, Madaus 1971). Very few examples of the use of questions like *"How"?* and *"Why"?* which through their requirement for recall and evaluation of background knowledge stimulate higher level skills of application, analysis, synthesis and evaluation were presented. It seemed that the emphasis in assessment, and probably therefore in teaching, had been laid on testable facts rather than on the syllabus aims. A clear need could be seen for an easily accessible, fair and accurate technique to evaluate this third of the Music curriculum more thoroughly without distorting the overall programme.

The Music courses for the new Tasmanian Certificate of Education (TCE), implementation of which began in 1990, contain broad objectives as did the previous syllabus, but evaluation is more strictly regulated through the use of compulsory assessment criteria. As each of these criteria must be evaluated thoroughly, there is a danger that the teaching of these courses will come to be dominated by them thus distorting the curriculum. Included among the objectives for the 9B and 10B Music courses are the following listening related items :

"an awareness of the need for accuracy and sensitivity in music"

"skill in listening in order to describe and make informed judgments"

"understanding ...rhythm, melody, texture, structure, idiom and style "

These objectives make it clear that the development of listening skills is important, yet only one of the assessment criteria makes any reference to listening :

"identifies and describes the main characteristics of music"

Other assessment criteria do require listening skills for satisfactory musical completion, but not all aspects in the objectives are even covered in this indirect fashion. An awareness of the need for *"accuracy and sensitivity in music"* could be developed through very careful teaching of solo and ensemble performance skills, and if carefully structured both performing and composing tasks could also teach and stimulate understanding of specific musical elements such as *"rhythm, melody, texture, structure, idiom and style "*. the evaluation of *"skill in listening in order to describe and make informed judgments"* is less easily placed in an active, practical context, and it is does not seem to be included in any of the assessment criteria. It could be that the lack of an available assessment technique for this objective, which has led to its omission from the assessment criteria.

The Evaluation of Music Listening

In the classroom the assessment of Music Listening has often been dealt with by the identification of facts about the music such as instruments, forms, styles and periods in one word answers. These answers reveal the student's ability to recall and recognise previously heard sounds, but do not show how the student arrived at their decision, or even whether they did recognise the sound or merely guessed its identity. Essay type answers and assignments are also commonly used, but they tend to rely heavily on memorised or researched background facts, so that they are knowledge based rather than evidence of actual listening experiences and student reactions to the music.

Active listening and music teaching text books often provide teaching material and classroom methods, and some also include assessment procedures, which teachers can use to ascertain how well the objectives of the method have been attained. *"Upbeat"* (Leask 1989)" has separate volumes for Levels 1-5, each of which deals well with the identification of the separate elements of music, and presents the complexities of each gradually in an endeavour to widen students' perceptions of Music. The value of objective evaluation in

the music classroom is stressed, as it can not only be an equitable method of assessing the progress of every student fairly, but can also assist in the identification of the musically talented student, in providing diagnostic evidence, and in ascertaining the success or otherwise of the teaching programme. Every unit focusses on a separate musical element and has a formal evaluation section using Listening, Performing and Inventing (composing) activities centred around that element. Both the teaching and evaluation components of each unit make extensive use of recognition and comparison activities in a practical context in an effort to retain the totality of the music experience.

Similarly *"Discover Music Making"* (Stowasser 1989) also makes extensive use of practical work in each of the three musical activity areas to teach students understanding of musical concepts, such as rhythm, form and texture. Each unit is based around a style of music rather than an element, with activities suitable for students with various levels of previous musical experience. The text is aimed at the senior secondary level and although no mention is made of evaluation, there seems to be an expectation that students will become immersed in tasks involving a synthesis of listening, composing and performing and that satisfactory completion of a task would indicate an integrated musical achievement. Evidence of an individual students' proficiency in each of the areas except performance would be, however, hard to establish in the practical situation as this is usually a cooperative situation.

As will be shown in the next chapter, it does appear possible to test for recognition and probably understanding of the separate musical elements through both musical and artificial audio tests, but are these tests relevant? Music is always experienced as a whole, a gestalt combination of musical elements, and it is almost impossible to isolate any one element in a true musical performance. Pieces containing only one element are rare. Even drum kit solos which appear to be pure rhythm, also contain volume in their climaxes (often an effect created by faster notes in rolls) and timbre as the drummer shifts from snare to tom tom to bass, changes sticks and brushes, incorporates hi hat, crash and ride cymbals, or

merely moves from one area of the head to another. It is therefore debatable how far findings based on the examination of one element in isolation can be applied to normal musical activity.

In addition, music only exists in time as a passing aural sensation. It is therefore almost impossible to determine exactly what has been heard at any one moment. It is also not clear whether it is possible to evaluate the entire listening experience. Investigations are currently being continued into the nature of the listening experience, how it is perceived by and affects the brain and the body, and what specific involuntary responses are caused by it (Roehmann and Wilson 1988).

As every listener also brings memories and prejudices derived from their previous experiences to each fresh listening experience, they may observe the music from different perspectives and in fact notice and ignore different aspects of it. There does seem to be, for instance, some evidence (Hufstadter 1977) that very young children have not yet developed the ability to perceive the harmonic elements of music, whereas this is one of the prime concerns and often the first aspect noticed by jazz musicians. There does seem to be a problem in determining exactly what an individual hears and therefore in establishing what has been tested.

The validity of testing a musical experience by a verbally based one can also be challenged on various grounds. How is music experienced by the brain? Does the brain respond in a totally iconic way or does language play a part in response to music? Does the act of responding to a test procedure alter the actual listening experience? and what about the time lag between the acts of listening and responding? These problems are all part of the normal experience of listening, and should not be regarded as insurmountable. Perhaps a suitable way to approach this problem would be by looking at the responses to music given by those who are acknowledged to be good listeners.

At a concert, two experienced musicians may hear and notice all the elements that make up the music but focus their listening

attention and discussion of the performance quite legitimately on differents aspects of the music, one noting the technical mastery of the soloist and the other appreciating the strict attention to stylistic detail and performance practice. Each person listens from their own conceptual background and has their attention slanted towards those aspects of music which they find the most interesting, satisfying or important in that performance. If they were both equally aware of all the details of the music, even though they did not think it necessary to mention them, it could not be said that one was a better listener than the other, but that they just had different reactions to it.

Perhaps the problem that should be addressed is not WHAT do they hear? but WHY do they focus on specific elements within the music? A skilled listener will assume awareness of the basics such as speed and volume, and not deign to mention them, focussing their attention on more complex concepts such as counterpoint, orchestration, form and performance practice, whose understanding requires a high level of awareness of the simpler more basic elements. An analysis of why the listener isolated some elements rather than others, and what justification they gave for their judgments, would perhaps give a more accurate estimation of their level of listening than a mere quantification of the elements they consciously mention. An evaluation of the quality of responses, rather than the quantity of aspects recalled, could also be a more effective method of evaluating those curriculum aims which deal with understanding and making judgments.

Summary

Performing, Composing and Listening are the three main musical skills, with listening being an integral component of the other two skill areas, not merely an audience skill. Of the three, Listening is also the musical skill most commonly practised in the Western world, and a vital part of musical education in all cultures.

Certificated school Music courses acknowledge the importance of each of the skill areas and require assessment in all three. In Tasmania, evaluation procedures have been developed for Performance and Composition, but procedures presented at Moderation meetings for the evaluation of the syllabus components concerned with Listening seem to be incomplete. A technique is needed which can not only be an acceptable part of summative assessment, but which can also identify those musical elements not fully understood so that reteaching and relearning can take place. It needs to be a procedure which is easily accessible, fair and accurate, and capable of being applied without requiring extensive tuition that might distort the overall Music programme. It would be useful if it were also capable of leading students into an awarenesss of how musical elements combine to create the sound effect and therefore how the emotional aspects of the music are created.

Aims of this Study

In this study an attempt will be made to develop an evaluation technique for assessing responses stimulated by listening to music, which can be used for music in the classroom situation. Principles evolved by previous research will be used to devise a test format that can be altered to suit other situations and pieces of music. It is intended that it should be an evaluation technique that makes sense to students, and is simple and quick to use.

The following chapters will show how these aims were carried out. Chapters Two and Three will outline the search for a suitable existing test or a theory to use as the basis for a Music Listening test, which led to the decision to use the SOLO Taxonomy. The development of the connection between SOLO and Music and its investigation in the Pilot Study will occupy Chapter Four, and the development and usage of the SOLO test will be described in the Main Study in Chapter Five. Chapters Six and Seven will be devoted to an analysis of the data from the Main Study and an examination of the reliability, validity and useability of the technique, and conclusions as to the value of this study, the implications to be drawn from it and directions for future work will be contained in the final chapter.

Chapter Two

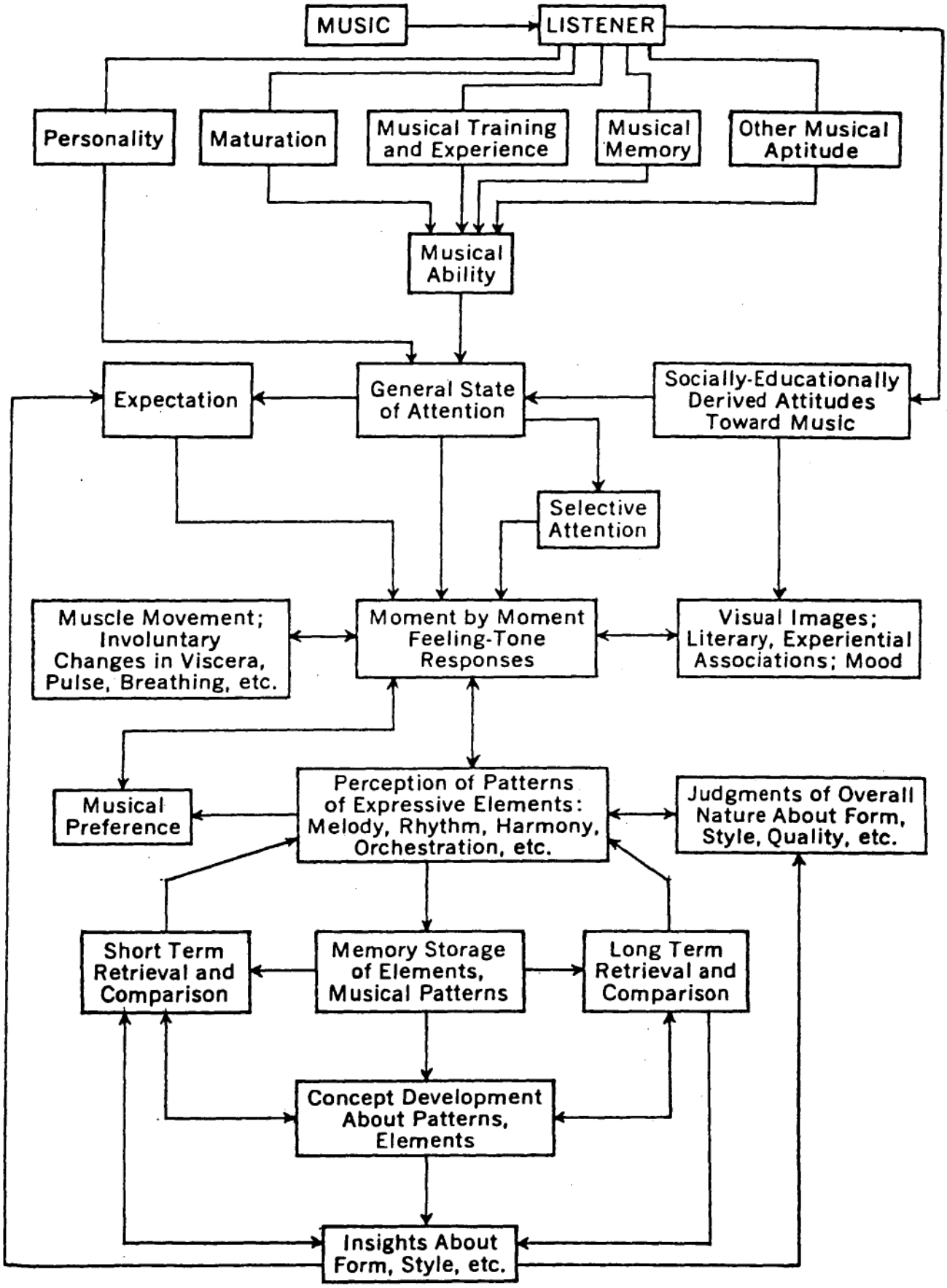
Review of Literature on Music Listening

Introduction

As the aim of this study is to evaluate student responses in the area of listening to music, this chapter will focus on thinking and research which has been carried out in that context. There have been two main approaches, which have been admirably summarized by Gardner (1984) as the "bottom-up" and "top-down" approaches. In the former, each element of music, for instance rhythm or pitch, is presented and tested separately. In the "top-down" approach the entire Music Listening experience is studied by examining responses to actual pieces of music. A third approach, the "middle-ground", has also appeared, and this uses short extracts of music to present a genuinely musical experience, whilst still allowing accurate identification of what has actually been heard. No matter which approach is used however, the listener's responses are always influenced by factors other than the music being presented, and these external influences upon the listener's response must always be borne in mind.

Figure 2.01 Influences upon the Listening Experience

from Prince (1972)



Influences upon the Listener

Influences upon the listener are many and varied, as can be seen in Figure 2.01 (Prince 1972), which was drawn up to place completed research projects and theories on Music Listening into perspective. Many variables are brought to each listening experience, and these govern how the stimulus is perceived at that particular moment, how it is responded to, how it is remembered, what is learned from it and how it is later used to influence responses to future listening experiences. Personality, maturation, the attitude of the social culture towards music, musical training, experience, memory, aptitude and ability are all seen as important determinants in the amount and type of attention which the listener gives to the musical stimulus, and they are all therefore important influences on responses.

Prince's variable factors fall into two groups : musical and psychological. Training, experience, knowledge, musical memory, aptitude, ability and preference are all basically musical factors; whereas personality, maturation, state of attention, attitude towards music, and expectation, could all be classified as more psychologically based factors. These variable factors are each in their turn influenced and altered by every fresh listening experience. Responses are also dependent on the musical character of the stimulus, the social background, the ambience within which it is presented, and whether it is a live or recorded performance in public or private.

Other commentators on factors influencing listener responses have focussed on different areas. Farnsworth (1969) noted that personality, mood preceding the experience, attitude towards music, the particular piece being listened to and the meanings of any lyrics were important; whereas Hopkins (1979) considered attention and willingness to listen as being essential for successful listening. Social class and mental age were influences on responses in tests carried out by McDonald (1974), whilst West, Howell and Cross (1985) noted that their subject response patterns were

affected and their musical experiences were controlled by their cultural histories. Previous musical experiences were also found to be factors affecting responses in research by Conley (1981), who also found a relationship between the perception of musical complexity and musical background. Experiences playing musical instruments and age level were listed as influences by Nierman (1983), whilst the role played by expectation and variation was added by Simon and Wohlwill (1968).

Many of these influences on the listening experience may be unnoticed by the subject and there are also involuntary physical aspects of listener responses of which the listener may also be unaware. These have been put at the heart of Prince's diagram and consist of changes in bodily functions such as breathing and pulse rate, emotional and associative feelings, in addition to the ability to hear sounds.

Many attempts have been made to determine what happens inside the body and brain in response to music. Fascinating studies of the way prenatal babies respond to music (Shetler 1990) and how birds perceive music and learn to sing (Hulse 1990, Nottebohm 1980) have been carried out and many complicated pieces of audio, medical and neurological equipment have been devised to give precise measurements of physical reactions to music as in recent experiments by Crummer, Hantz, Chuang, Walton and Frisina (1988).

Whilst monitoring the brain's production of alphawaves, Wagner and Harding (1987) confirmed that, as reported by other authorities (Wilson 1985), electrical activity in the right hemisphere of the brain was stimulated by music. The triune brain theory and music therapy experiences also suggest that initial responses to music may be at a *"nonverbal, non-analytical, yet powerful"* level (Harvey 1986), so that these apparently non-musical responses may in fact be the deepest and most natural way for humans to respond to music.

An inventory of responses to music has been put forward by Arthur Harvey (1986), an expert on music and the brain. In this inventory he lists four categories of response : physical, affective, transpersonal and cognitive. The physical category involves the physical responses controlled by the limbic system in what is considered to be the primeval area of the brain, and these are *"pulse rate, blood flow and pressure, breathing rate, muscle tone, pupil dilation, glandular activation, metabolic rate and neural activity"*, which can all be altered by changes in the music. Affective responses occur as a reaction to the feelings symbolised in the music, and the resultant mood changes controlled by the limbic system. The transpersonal response *"is one that many individuals experience, but often do not discuss with others, an imagery-evoking or consciousness-altering experience"*; and the cognitive approach involves awareness and analysis of the musical elements of the piece. Physical, affective and transpersonal responses could be categorised by these theorists as being basically right brain dominated experiences, with the cognitive response being thought of as being governed by the left hemisphere of the brain.

Cognitive responses and their interrelationship with each other and the listening experience are shown clearly in the lower section of Prince's diagram (Figure 2.01). Once again the role of previous musical experiences in the formation of thoughts about music is seen as essential and unavoidable. Storage, retrieval and comparison with music previously heard through use of the memory affect recognition of both musical patterns, such as rhythm, and the growth and development of musical concepts, such as form. The listening experience itself is also shown to be influenced by preferences towards particular styles of music, though this may perhaps be thought of as a factor which may inhibit the development of new ideas as it is not shown as leading towards fresh thoughts on the diagram. This may be however be due to preference being regarded as an affective response rather than one which is objective or cognitive. Recall of past experiences however is shown here as leading toward fresh insights and judgments.

Research Investigating the Separate Elements of Music

As music education has been seen mainly to deal with the cognitive aspects of music, it is this more objective, analytical side of the listening experience which attracted the earliest and most intensive research interest, and has led to the innumerable studies of the separate elements of music such as pitch, timbre, tempo, rhythm and harmony. Investigations into the ability to recognise and discriminate between the separate musical elements, "the bottom up" approach, began at the end of the nineteenth century, and have been continued since with many, mainly unpublished investigations. Attempts to relate Piagetian developmental stages, and in particular his theory of conservation, to the perception of music, have led to many studies involving students of differing ages, and to a growing desire for greater knowledge of the ages at which awareness of separate elements emerges, which would lead to the creation of an overall developmental picture can be created upon which a systematic music ducation curriculum could be based. The theories generated by research in this area have also led to the development of numerous commercially available standardised tests for musical aptitude, ability and achievement.

Most of the standardised tests include pitch tests involving two tones which must be rated as higher or lower, up or down, or same or different. The results of this type of task can be misleading, as Andrews and Madeira (1977) showed in their study which tested the hypothesis that the language used to frame questions could affect responses. Students aged 6-8$^{1}2$ years were tested with a variety of visual and oral tests, which assessed the ability to discriminate between pitches an octave apart and the understanding of the terms *"high, low; higher"* and *"lower"* and the use of these terms with the same pitches. In the first test the higher note was associated with a small pig and the lower note with a large pig, and as each note was played subjects moved the respective pigs towards their barns. Students performed well on this test, but gave fewer correct responses when asked to identify the

same sounds as high or low, or to identify the second of a pair of sounds as higher or lower. The significant difference between scores on these tests was attributed to the child's inability to deal with relational language, but it could also be because the first test only required the sounds to be matched to the appropriate pigs, whereas the other question required two steps to be undergone, identification of the sound and the establishment of its pitch relationship. Obviously great care should be taken when setting tasks and framing questions, and tests involving visual cues similar to those in this test have frequently been used with very young children.

Later studies using short melodic fragments to investigate tonal memory have taken great care to avoid any possible language biases. Work with young children carried out by Edworthy (1985), and Pick, Palmer, Hennessy, Unze, Jones and Richardson (1988) has shown that they are capable of remembering melodic patterns and detecting transpositions, and that despite transpositions they can still identify tunes and distinguish between alterations of the same tune. Experiments with normal and mentally defective students carried out by Zenatti (1976) also showed that for both types of students tonal melodies were easier to handle than atonal ones, and that these abilities developed at about the same mental age.

The ability to detect changes in pitch seems to develop early, and this could possibly be accounted for by its being a primitive survival skill. Primitive man would also have needed to distinguish between the sounds made by different types of animals, predators in particular, and this could account for the early development of the ability to discriminate between timbres.

Jetter (1978), and Wooderson and Small (1981), have shown that recognition of the timbre of orchestral instruments seems to develop far earlier than is normally realised. In their studies children aged 4-6 were shown to have the ability to recognise sounds well. Lower frequencies, which give lower pitched sounds seemed to be recognised quicker than higher pitched ones, and some instrumental timbres, especially in woodwind instruments, alter according to

volume, which makes recognition difficult. The attack of the sound, pitch and volume were also shown to be factors affecting recognition of timbres. As attack and decay are features which give character to individual sounds, this was not an unexpected finding.

The perception of speed has also been investigated through experiments in the identification and estimation of changes in tempo carried out by Kuhn (1974) and Madsen (1979). These studies have shown that slower speeds are identified better than faster ones. In addition, when estimating speed, tempi are generally underestimated, but musicians seem to be more accurate at this than non musicians.

Tempo and rhythm are inextricable aspects of a piece of music, contributing greatly to its mood and character, but in this type of research they have been separated. Regular and irregular rhythmic patterns were studied by Sturges and Martin (1974) in a study to determine which were the more memorable. Regular patterns were found to be easier to remember than irregular ones. Heyduk (1975) also found that complex patterns like syncopation were preferred to more straightforward ones. It seems that the pop music industry is well aware of both these points, as most of the music in the Top 10 consists of a very regular and straightforward rhythmic basis laid down and repeated by drums, bass and rhythm guitar or keyboard, with more complex syncopations in the tune and fill-ins working against this. Heyduk also investigated preference for textures and found a clear bias towards pieces which were more complex in both rhythm and texture, a finding supported by Radocy (1982), whose research showed a strong connection between preference and complexity.

Complexity in music is often expressed through variations to an initial harmonic structure, and the ability to discriminate between harmonies, as it involves a longer term use of memory, may be a more advanced aural skill. Farnsworth (1969) supports this theory and states that harmonic discrimination is a learned skill, influenced by several factors : return to the key tone, interval sizes, and the movement of the voice parts. However in investigations with children from Kindergarten to Grade 3, although an improvement in

the ability to discriminate was shown to correlate with age, the ability to discriminate between tune harmonisations, Bridges (1965), and between chords, Hair (1977), was present, even at this very early age.

Developmental Stages in Recognition of Musical Elements

Knowledge of the ages at which aural discriminations develop is essential for the development of music teaching curricula, so research comparing the aural recognition of several musical elements over various age levels has also been widely undertaken. Even though each study may have been concerned with one or a few aspects of music, the results of these have led to some suggestions of a developmental sequence of aural abilities being put forward.

McDonald (1974) working only with Grade 4 children, reported that they found it easiest to discriminate between degrees of loudness, followed by tempo, and then pitch. The students from Grades 1, 3 and 5 who were tested in O'Hearn's study (1984) found timbre easiest, then dynamics, pitch and duration. Hufstader (1977), working with students from Grade 1-8 found that discrimination of timbre appeared at Grade 1, rhythm and melody at Grade 5 and harmony at Grade 7. He did however also point out that, though this sequence was true for the group overall, it did not hold for all individuals within the group. Hedden (1981) in another study added that low socioeconomic and culturally deprived groups scored lower on tests, and he also found a definite developmental sequence of musical awareness that began with volume, and continued through tone colour, melody and rhythm to the eventual awareness of harmony. Most researchers report that regardless of when recognition of an element first appears, the ability to discriminate aspects of every element is better in older students.

Standardised Tests of Musical Ability and Achievement

The research studies dealing with these isolated elements of music are always carefully carried out especially with regard to experimental design, use of controls and analysis, and are fascinating to read. It is however questionable whether or not their findings are relevant to, or valid in, more normal musical situations, where pieces of music are encountered which contain all elements. These studies have, however, led to the development of standardised tests, many of which are still commercially available. Many of these tests are commonly used to determine entry into instrumental programmes in high schools and into music courses at tertiary level, especially in the USA. In Southern Tasmania entry of students at Grade 5 into the Education Department's Primary School Instrumental Programmes have for some time been partially determined by use of the Bentley Musical Ability Test or the Selmer Music Guidance Survey.

The Bentley Measure of Musical Ability is a relatively short test, suitable for ages 7 - 12. It consists of four tests - Pitch, Tunes, Chords and Rhythm - which test both aural discrimination and short term memory. All the sounds used in the pitch test are produced by an oscillator, and as 80% of them are movements of less than a semitone (the smallest distance used in Western music), it is a valid aural perception test, though its relevance to actual musical situations and therefore its use as a predictor of success in Western music could be challenged.

The Selmer Music Guidance Survey is a similar four part test, with the sounds usually presented on a tape played with a synthesizer sound. The first, Pitch Recognition, consists of pairs of isolated notes for which the student must state whether the second note is higher, lower or the same as the first one. No bass notes are included as the note range is from G to F^{111}, and the intervals are those commonly found in Western music within the range of a minor 3rd. In Chord Memory, the second test, the students hears two chords and has to identify them as the same or different. This is quite a difficult test as all the chords are fairly high pitched

and the differences are sometimes quite subtle. Same or different is also the response required in the third test, Melody Memory, where pairs of four note melodic fragments must be compared, once again with no low pitched phrases included, however as most of the melodic patterns could be found in Western music, this is a relatively musical section. Rhythm Memory, the final section, involves a same or different comparison between one or two bar rhythmic patterns.

The Gordon Musical Abilities Profile (M.A.P.) is a typical aptitude test. Its design was intended to minimize musical achievement so that the most basic factors of musical aptitude- musical expression, aural perception and kinesthetic musical feeling- could be expressed. Subjects listen to three taped programmes- "Tonal Imagery, Rhythm Imagery and Musical Sensitivity"- and make judgements based on what they hear, either identifying altered notes in repeated phrases, or deciding which of two versions of a short phrase they prefer. Both aural exactitude and culturally based preferences for phrasing, balance and style are thus included. The test is obviously culturally biased towards the conventions of Western music, and, as it is a purely aural test involving checking off lists, it could be debated whether or not it really does achieve all its stated aims.

The three tests outlined above are typical examples of the collections of short tests found in the commercially available standardised tests, which also often use taped synthesized sounds as the aural stimuli, and test pitch with clinical accuracy. It has been stated (Boyle and Radocy 1982) that *"Musical ability is best measured by an assessment of realistic musical skills, intellectual ability, academic achievment, environmental stimulation and ... physical attributes"*. As the presently available tests do not cover all these areas their use as predictors of student achievement in music is therefore questionable.

Some tests of musical achievement do try to avoid the artificial situation caused by the use of synthesized sounds, and although they may use taped sections they usually include other types of test as well.

The Watkins Farnum Performance Scale (1954) and the Farnum String Scale (1969) for instance, suggest that the tester play the tests on instruments appropriate to the purpose for which the test is being used. For band programmes it is suggested that the tests be played on a clarinet, and for orchestral programmes a violin.

The Kwalwasser-Ruch Test of Musical Accomplishment for Grades 4 through 12 (1927, 1952) includes tests using standard music notation, where the student is required to select and identify the correct musical symbols. The Colwell Music Achievement Tests (M.A.T.), rated as suitable for Grades 3 to High School, are aurally based with four sections : pitch discrimination; auditory-visual and tonality; tonal memory, pitch and instrument recognition; and style, texture, rhythm and chords.

The focus in these tests on what seem to be isolated musical elements in non musical contexts is disturbing, and further problems emerged in the analysis of music perception tests for item difficulty carried out by Richard Colwell (1987). He sorted the items from these tests into four groups according to degree of difficulty and found that most tests used the easier items. Easy items included : recognition of the direction of intervals, identification of individual instruments, lengths of notes, volume, melodic recognition, counting the number of phrases, deciding whether or not the final cadence is finished, and deciding same or different for melodies, chords, rhythms and textures. Naming intervals, recognising instrumental families, determining the correct balance, finding the melody in a simple piece and deciding if the texture is homophonic or polyphonic were rated as harder items. The more difficult items were : identifying the tonality as major or minor, counting the number of parts, identifying the chords, stating the relationship between the rhythm and the beat, identifying the type of accompaniment and defining its role, recognising the type of melodic movement, recognising cadences, matching sounds to notation and given "doh", singing a specified pitch. The most difficult items were those requiring items of academic knowledge to be identified and then integrated, such as

explaining the function and movement of chords. He came to the conclusion that the emphasis on the easier (and perhaps less musically relevant) items was because these tests could be used with a wider population and were therefore more commercially viable.

The Australian Test for Advanced Music Studies (ATAMS) compiled by Bridges for the ACER in 1974, is one which attempts to resolve these problems. The test is divided into three sections : "Tonal and Rhythm Memory and Musical Perception", "Aural/Visual Discrimination, Score Reading and Understanding of Notation" and Comprehension and Application of Learned Musical Material". Students are required to listen to extracts of "real music" and to demonstrate various levels of skill in relating the sounds heard to music notation.

Many researchers have used standardised tests in addition to their own specific ones, often to establish a reliability rating, and higher scores have generally been found amongst the higher socioeconomic groups. For instance Hill (1968) used the Gordon M.A.P. with children from Kindergarten to Grade 6 and found that the higher scores came from those in the higher socioeconomic background. This finding was replicated by Swickard (1971) who used the Colwell M.A.T. with a sample of 3000 students from Grades 4-6. However Standifer (1970) found the reverse effect in a study involving a new teaching course, but he attributed this result to the intense interest and pride in this class which was generated within the school and the outside community, which led to a new sense of excitement, worth and purpose amongst the students in the lower socioeconomic school, and to increased motivation in the class involved in the study.

Standardised tests are often used to select students for special courses or instrumental tuition, but they do not seem to be used diagnostically in the classroom. These tests are often good at isolating and examining particular elements of music, but they do not usually deal with real music, which is the material used in the music classroom. Even when teaching the recognition of instrumental timbres, teachers use recordings of melodies selected from the standard repertoire rather than using recordings of isolated notes.

The "top-down" research approach using complete pieces of music or sizeable extracts seems therefore to be more relevant to the normal school situation. It does however carry with it the difficulty of establishing exactly which segment of music the subject is attending to, and responding to; so studies using the "top-down" approach are subject to more uncontrollable variables.

Research Investigating Responses to Pieces of Music

The earliest researchers in the 1920s and 1930s, interested in music appreciation and aesthetic responses to music, collected responses which they used to classify their subjects into types of listeners. Later studies used these statements about types of listeners to develop the lists of adjectives and phrases, referring to the music which they had just heard, from which their subjects were required to choose. Had the responses always fallen into the same categories or types, conclusions would be clear, but as fresh research methods appeared so did new types of response requiring fresh categorisation.

Early researchers concerned with Music Appreciation wanted to find effective methods of teaching it. In 1927 Ortmann reported on studies investigating types of listener responses. The response types which he observed were the sensorial, the perceptual, and the imaginal. The sensorial response is stimulated by the *"raw sensory material"* of the music, which he defines as the pitch, intensity, duration and 'quality'. This is the untrained reaction which predominates in listening to popular music and the light classics. The perceptual type of response is essentially an interpretation of the sensorial effect based upon the awareness of the relationships between aspects of the purely sensorial elements such as phrases, contrasts, melodic outlines. The imaginal response is a representational one which results from the ability to anticipate what is yet to come and to make judgments about the music based on the combination of recall and anticipation.

In a study with 15 adults Myers (1927) found four types of listener response : intra-subjective, being the sensory, emotional or conative experience aroused by the music; associative being reminders of other experiences both musical and non-musical; objective being the purely musical use or value of the music, a critical or analytical attitude; and character, being the personification of music with human characteristics.

A much later study by Yingling (1962) classified responses into similar categories. Groups of college students were asked to listen to the same pieces of music and then answer the question *"What does this music mean to you or what does it do to you or for you ?"* Responses were originally classified into extra musical and musical categories, but eventually four categories were used : associative (extra musical associations), emotional (reponse involving own emotions), intellectual ("engaging the intellect", which includes Ortmann's perceptual and Myers' objective categories) and sensory (reactions involving personal movement or tension). Yingling was not concerned with the musical validity of the responses he received, merely with classifying them. It could be that the nature of the question he asked, with its emphasis on a personal reaction, influenced the type of responses given, and therefore the categories into which they could be classified.

Other researchers have been more concerned with their subjects' abilities to identify what they hear in the music, trying to ascertain their objective rather than subjective response, and many ingenious testing methods have been developed. Probably due to the need to simplify the analysis of responses, and eliminate or reduce the bias towards the more linguistically fluent subjects, most researchers have given their subjects checklists or rating scales instead of the opportunity to express themselves freely either verbally or on paper.

Kate Hevner (1935) devised a clock face set of grouped adjectives, which provided a resource for subjects to select from when describing the music. This has also been used as a starting point by other researchers, and was revised by Farnsworth (1954).

Various versions of this technique using adjectives, descriptive phrases or statements have been used effectively to assess both extracts and complete works by Standifer (1970), Nierman (1983), Olson (1984) and Asmus (1985). Valovy (1981) even gave primary school students a choice of fairytale characters to match the music. Later when studying music appreciation, Hevner (1956) used a different technique, a free response plus a questionnare with 42 statements to be rated on a 5 point agree-disagree scale.

In a study reported in 1980 Payne used sets of statements classified into six categories: aesthetic emotional character, human emotional subject matter, form or texture, historical significance, orchestration, and extra musical implication. She found that the primary appeal of music was an emotional one, though this effect was less strong with trained musicians and increased familiarity with the music. She postulated that training gives understanding, which leads to new interests within the music and fresh insights upon repeated hearings.

Olson's Measurement of Musical Awareness (1984) incorporated feeling, analyzing, and judging, in the study's definition of musical awareness, and aimed to gauge awareness of music at three levels : musical details, associations and mental images, and accentuation. [Accentuation was here taken to refer to the depth of involvement with the music felt by the subject, and was considered to be *"a unifying quality which brings together objective and emotional perception"*.] Measurement of the levels of awareness was carried out by means of an extensive questionnaire which subjects completed after hearing a piece of music. Subjects were required to make a choice from a selection of phrases provided, as for instance in the section devoted to Musical Details, where Form had to be identified as either *"main melody used in various ways"*, or *"unity exists through repetition"*, or *"larger form develops from single melodic ideas or motives"*. Scores on this test were found to correlate with composite grades obtained for the subjects on music theory and were significant at the .01 level of probability, so it was considered that the Measurement of Musical Awareness could be considered an achievement rather than an aptitude test, and as such

could be suitable for diagnostic use. Olson also felt that the test did measure musical awareness. Information on the subject's musical environment was obtained by a pre-test questionnaire, and activities which involved intensive listening appeared to be major factors in developing musical awareness, whilst repetition seemed to be a factor in increasing it.

Crickmore's (1968) interpretation of Musical Awareness examined enjoyment of music rather than traditional music appreciation. He used ratings on a three point scale, which he called a Syndrome Test, to measure each subject's degree of involvement with the music. He also asked subjects to add a brief comment about the music and their involvement with it whilst listening. His subjects, who were mostly engineering students, gave reactions which when compared to independent measures for personality, intelligence and musical intelligence, showed no relationship with his measure of musical awareness, so he drew the conclusion that his test was measuring musical awareness.

Nelson (1984) worked with violinists aged 3 - $16\frac{1}{2}$, who were taught two tunes and then asked questions relating to their opinions of the pieces both as works of art and as part of their repertoire. He found that age was the main factor in determining types of responses. Younger students were more egocentric, developing an awareness of the musical elements within the pieces later, and coming to appreciate the complexities when older. He also felt that "a more openended series of questions is needed...to give each child a sufficient opportunity to express views and opinions."

In 1957 Copland, a leading American composer and musician, formulated a view of listening as an activity on three planes, which seems to draw together ideas from both researchers and musicians, which seems to draw together ideas from both researchers and musicians and has been accepted by many musicians. Copland's proposition is that music can be heard on three separate levels, called planes, which form a hierarchy of responses to music. The

initial response is on the sensuous plane, with higher level responses moving through an expressive plane, into the plane at which musicians listen, the sheerly musical plane.

The simplest is the sensuous plane, listening for "the sheer pleasure of the musical sound itself", *"absent-mindedly"* basking in the sound. This does not denigrate the value of the choices of sound qualities (timbre) made by the composer but stresses that *"his usage of sound forms is an integral part of his style and must be taken into account when listening"*, not regarded as the whole.

The expressive plane involves those who like to imagine that music has moods and meanings, or tells stories. It involves associations outside the purely musical realm. This plane is somewhat controversial amongst musicians as not all composers wish to acknowledge the expressive qualities of their music (unless it is descriptive or programme music) and the issue of music as a device of communication is constantly debated. Copland believes that *"all music has a certain meaning behind the notes"* which *"constitutes...what the piece is about"*, but he does admit the difficulty in defining that meaning in detail, and denies that it is an essential component of a piece of music, and as musical moods can be changed with interpretation and the listener's attitude, the meaning can never be fixed. *"Music whose meaning is slightly different with each hearing has a greater chance of remaining alive"*, and it is perhaps this aspect of them that makes the greatest compositions and the most beautiful tunes the hardest to describe satisfactorily. No matter how much musicians argue as to whether or not music can act as a direct communication medium, reactions on this expressive plane exist and are often amongst the first reactions for nonmusicians, as even the earliest research in this area (Ortmann, Myers, 1927 cited above) has shown.

Copland's third and highest plane is the sheerly musical one, and it is to this plane that most of the literature on music is directed. The listener is encouraged to hear the music from the same viewpoint as the composer, being actively aware of melody, rhythm, harmony, timbre, texture and formal structure, and conscious of all

the possible avenues the composer can select as the music progresses, *"carried away by it yet coldly critical of it. A subjective and objective attitude is implied in both creating and listening to music"*.

Copland has put forward a developmental sequence of listening responses, beginning with the sensuous plane and progressing through the expressive to the sheerly musical plane. He regards the sheerly musical plane as the most important, and from the composer's viewpoint this is undoubtedly so; but most music researchers have classified their subjects' responses without making value judgments about the categories.

The "bottom up" approach to Music Listening splits musical experiences into their smallest identifiable components and can thus be precise about what has been heard, whereas the "top down" approach using complete pieces of music or sizable extracts, cannot identify the musical moment the subjects respond to with any degree of accuracy. A "middle ground" approach has recently emerged, which by using small segments of music with clearly identifiable aspects, attempts to control variables, whilst still presenting genuine musical experiences. Comparison of styles and discrimination of changes, in composed or altered musical extracts, are both tasks which can require subjects to identify exactly what it is that has been the basis for their judgments. If the researcher selects the material used carefully enough to reduce the number of options available, it is sometimes possible to find out what listeners have focussed their attention on, and what they have ignored, if not exactly what they heard at any specific moment.

Due, in part, to an interest in the possible application of Piaget's theories of conservation to Music (Serafine 1980, Larsen and Boody 1971) there has recently been an emphasis on young children and their ability to recognise changes in music. Some researchers investigating perception of changes in music with very young and therefore less articulate children (K-6), have tried to reduce the verbal bias by using visual cues. Hair (1981) devised a test in which a wellknown children's song was altered in 10 ways and

students were asked to say what had changed. In a later (1987) study she first elicited the verbal response, then two weeks later retested with visual cues, finding that the visual response was more accurate. McDonald (1974) and Simons (1976) both also used pictures to represent alternative aspects of the changes for responses from children K-4, and Simons found responses to be influenced by maturation and experience.

Perhaps the only researchers to have achieved a near normal situation, were Flowers (1983) and Herberger (1983) who allowed their subjects to hear the music in its entirety and and then express themselves freely. Both used a system of numbering sections within the music to identify the exact musical stimulus being described, and were concerned with the identification and explanation of changes within the music. Flowers' technique was to play the music through headphones, and as the subjects detected a change in the music they spoke a number, which was recorded plus the music onto a second tape. Subjects then replayed the second tape and wrote a description of the changes they had heard at each number. On the post-test most subjects detected fewer changes. Herberger's technique differed in that the control numbers were decided by the researcher and appeared both on a screen and on the lines of the answer sheet. Responses were written on the numbered lines, with some control numbers compulsory and some optional. This number technique allowed for the changes in moods, dynamics, impetus and form found in all "good" pieces of music to be noted. As Herberger's study also incorporated a formal appreciation-style lesson sequence between the pre- and post-tests, analysis of results could be on the basis of increases in number and quality of responses, and conclusions drawn from the results could have possibilities for implementation in classroom teaching practice.

The ability to identify and discriminate between small changes in music is an important factor in the appreciation of long pieces of music, and an essential component in the recognition of musical forms, styles and idioms. Recognition of musical styles is one of the aspects of music appreciation that concerns music teachers. It is a complex task, dependent upon the discrimination of many diverse

elements within the music. A composer's style is determined by factors such as their environment, historical setting, personal taste, craftsmanship and musical vocabulary, and when judging style it is almost impossible to detail all the cues used. Yet it seems from the research available that even young children can recognise styles well.

Students aged 6-19 were tested by Gardner (1973) with matched and unmatched pairs of carefully selected extracts. He was careful to remove any obvious cues such as orchestration, so that decisions had to take a variety of factors into account, and he found that despite their inability to verbalise their reasons, the 6 and 8 year olds could match styles quite well. There also seemed to be a developmental sequence in the types of verbalised reasons, which might be related to vocabulary and/or conceptual developments. The 6 year olds tended to give simple musical ideas (high/low, loud/soft, fast/slow) to show that the pieces were continuous, whereas 8 year olds had some awareness of the overall character of the music, which they explained through the use of nonmusical images relating to their own experiences by using words such as *"peppy", "churchy", "horse race".* By the age of 11, students had become aware of several variables within the music such as instrumentation, rhythm and texture and could accept the possibility that the extracts could be a discontinous selection from the same piece. They also listened more attentively, had little difficulty retaining the music in their memory, and tended to accept the music in its own terms rather than try to relate it to their own experiences. Students aged 14-19 had much more musical knowledge and used musical terminology. They *"had a sense of a musical piece as a structured entity which possesses continuities, direction, possibilities and implications,"* and could consider a large number of variables and their interrelationships. They were also aware of composers' varied usages of thematic and orchestration devices, and found it harder to make decisions because of their speculations.

A similar task was also set by Kate Castell (1982), who played paired extracts of both pop and classical music. She asked students aged 8-11 to imagine that they could see the musicians playing in a

room, then that they had left the room and coming back in heard and saw the second piece. The students had to say whether it was the same group of people playing the same piece or not. When asked for the reasons behind their decisions, she discovered that the older students were more fluent, but that both age groups lacked an adequate vocabulary to describe the pop music extracts which she used. The reasons given for their decisions were classified in 7 categories: same sound, tempo, instrumentation, melody, associative, categories (styles) and other responses.

Non Western music was included in a wide variety of styles used in a study by Hargreaves (1982). The students tested were aged 7-15, and were also played paired extracts, with the same question as above to be answered. In an attempt to reduce the age-linguistic bias, these subjects were restricted to a one sentence answer, but older students overcame this by including more points and clauses per sentence. The responses in this study were categorised as: objective-analytic, categorical, affective, objective-global, and associative.

Even though music teachers may teach towards Copland's sheerly musical plane, these studies seem to provide evidence that listener responses are determined not only by their ability to perceive aspects of the music and their knowledge about it, but also by their emotional and associative reactions to specific pieces and types of music. Evidence of the existence of Copland's other planes, the sensuous and the expressive also seem to have been shown in this research. In some cases affective responses seem to be very strong, and they are often experienced when music associated with emotional events in life, such as adolescence, is played. This is in part an explanation of the development of subconscious musical preferences and prejudices and the important role they can play in a response to music.

Music Preferences

Early work in the area of music preferences, was undertaken as part of studies into musical abilities, and had a definite cultural bias with serious classical music given a higher value by researchers than other styles of music. Modern studies, usually trying to be less biased, although still often slanted towards Western style music, are striving to measure attitudes towards various types of music and through this determine which factors lead to initial acceptance and eventual enjoyment.

Getz (1966) and Prince (1972) identified lively tempo, clear cut driving rhythm, a conjunct diatonic melody with repetitions and a variety of dynamic levels as the main factors leading towards initial acceptance and therefore willingness to listen closely. They related dislike to an angular melody, dissonant harmonies, slow tempo and minor key. Orchestral music was also preferred to vocal, with girls preferring orchestral and boys electronic and progressive pop music. Getz also found that both sexes were biassed towards music played by instruments they themselves played, and Hargreaves and Castell (1987) found that students preferred familiar melodies. Repetition up to 6-8 hearings was found to increase liking, a finding duplicated by Bradley (1971), Bartlett (1973), Heyduk (1975)and Herberger (1987).

Four separate taped music channels were offered to Grade 5/6 and University students by Geringer (1982) and the time spent listening to each channel was recorded. Preferences were found to relate directly to amount of musical training, with college music students selecting the classical music channels more often than other college and primary school students.

Teacher approval was the treatment factor in experiments by Greer, Dorow and Wachhaus (1973) and Dorow (1977), which detected increases in preferences for the type of music taught under high approval conditions. Peer approval was also examined and found to be influential.

Steck and Machotka (1975) presented 80 compositions representing 16 levels of complexity in a study that showed preferences to be influenced by complexity, novelty and intensity. Familiarity and complexity as aspects of preference were also scrutinised by Radocy (1982) in his "Test of The Hedgehog", where he examined Walker's theory of preference for psychological complexity with reference to music. University students rated complexity, familiarity and preference for pieces of music heard on a five point scale. The ratings for complexity were consistent within subjects, and preference and familiarity were strongly linked.

Summary

Research has been carried out into music listening skills with several aims: to ascertain what can be heard, and how this awareness appears and develops; to distinguish between students; to determine the types of response which can be stimulated by music; and the influences upon these responses. Many projects have been connected with teaching methods and have also attempted to evaluate the effectiveness of these programmes. Some of these findings could be related to the classroom, but none of the studies examined has produced a test or assessment system that could be used to evaluate either a good listener or the listening aims of the Schools Board Music courses as mentioned in Chapter One.

"Top down" research has isolated musical elements and produced evidence of the gradual emergence of listening skills, which have influenced both the teaching and selection of music for young children. This research has also been used in the preparation of standardised tests of musical ability, which as measures of ability or aptitude are often used in course admission procedures. As classroom music teachers however usually deal with actual pieces of music, there seem to be more possibilities for application to the classroom in the "top down" and "middle ground" studies.

Researchers dealing with actual pieces of music attempt to create a more normal situation, and various ingenious solutions have been devised to identify exactly what the listener is reacting to. These seem to work well in the individual or laboratory situation, but most would be unsuitable for use with a large class. A test is needed to suit the classroom situation.

The most promising aspect of these studies has been the classification of responses stimulated by open questions and the conclusions drawn from these about possible stages in the development of thinking about music, which could lead to judgments about the musical quality of student responses. Thinking by musicians such as Copland (1957) and work by researchers such as Gardner (1973) and Nelson (1984) seem to show that there is a possibility of using a structural analysis of student responses to order responses in terms of conceptual or structural complexity.

Chapter Three

Review of Literature

The Structural Complexity of Student Thinking

This chapter examines methods of analysing responses in terms of their structural complexity. Complexity can be shown through several aspects of a response : through its use of language, through the nature of the concepts it employs, and through the structure of the response itself. It could be possible for one of these to be developed for use in the classroom evaluation of Music Listening.

Responses can be rated according to the technicality of the language employed. In this case the assumption is made that knowledge of a word presupposes understanding of the concept behind it, but this, as classroom teachers know, is not always the case. Students, especially young ones, will often use words learned from adults without understanding their meaning, and only a detailed examination of the context will reveal the problem. Evaluation methods therefore that rely on the subject's use of language can be misleading, and perhaps a better strategy would be to focus evaluation upon evidence of the understanding of concepts.

Understanding how children develop their concepts of music and the world around them has been the focus of much research and the major concern of many music educators and psychologists. Many

theories have been put forward which could perhaps be used as the basis of an evaluation technique. This chapter considers both the results of research into the appreciation of Art and Music, and ideas about the organisation of thought, in the search for a suitable theory upon which to build an evaluation technique.

Response Structures in Art and Music

Music educators and researchers have usually focussed their attention on the musical nature of responses and have classified them accordingly (Yingling 1962, Payne 1980, Hargreaves and Colman 1981, Castell 1982, Asmus 1985). These response classifications could be useful, but as they do not indicate the relative value of each classification, it is difficult to see how they could be used in an evaluation of the quality of a response. There are however also some studies which reveal a hierarchy of skills, and these might be more useful for evaluation as they set a standard against which other responses could be compared.

The study by Nelson (1984, cited in Chapter 2) with violin students, asked questions dealing with separate aspects of aesthetic judgment, rated the responses as egocentric or showing musical awareness. Although all responses were quite short, some were more complex and were awarded extra points. Although this study is extremely interesting, accepts the problem of rating the quality of musical responses, and attempts to provide a solution, it does not present sufficient rating categories to be useful in the secondary classroom situation where finer qualitative distinctions between responses and a mechanism for evaluating longer responses are needed.

Gardner's study of reasons for stylistic judgments about music (1973, also cited in chapter 2) goes further, showing how awareness of different aspects of music could be related to the students' age. The youngest students saw the pieces as being continuous, and used basic musical ideas, such as speed and volume, in their reasoning.

Personal imagery to show the character of the music was focussed on by the 8 year olds, whereas the 11 year olds were more musically objective mentioning more aspects of the music, and the oldest students (14-19) considered many variables and could understand their interrelationships. This trend in the type of responses seems to mirror Copland's concept of the three planes of listening, sensuous, expressive and the highest, sheerly musical (Copland 1957, also cited in Chapter 2). Once again these categories of response are helpful when sorting responses, but they do not deal with evaluation of quality within the the musical plane, which is the main one dealt with in the Schools Board syllabus. Studies in the area of Art Appreciation however, are primarily concerned with the recognition of quality and perhaps have some relevance to this problem in Music.

In an examination of responses to visual art Rosenstiegel, Morison, Silverman and Gardner (1978) made response classifications which when related to the ages of the students giving them, showed trends in the stress laid upon each aspect at different ages. In responses to questions based on six topics in art appreciation a study by Parsons, Johnston and Durham (1978) found that there were levels of quality in each of the topic areas. Further work reported by Parsons (1987) has led to the hypothesis that there are five stages in the development of Art Appreciation : Favouritism, Beauty and Realism, Expressiveness, Style and Form and Autonomy.

Each of these stages shows a deeper understanding of paintings than the previous stage, demonstrating a sequential development from dependence to autonomous judgments. The first stage, Favouritism, is where children demonstrate *"intuitive delight"* and universal acceptance of all they see, and base their value judgments upon subjective associations. Beauty and Realism, is a stage focussed around the subject of the painting, where value lies in the accuracy of representation; whereas the emotional experience created by viewing the painting is the essence of the Expressive stage. In the next stage, Style and Form, the main concern is with the compositional layout, media and techniques used to create this expressive effect. Appreciation of the artist's craftsmanship is

developing at this stage, and culminates in the final stage, Autonomy, where conventions are questioned and judgments are made on the basis of personal interpretations of them. These five stages can be seen to bear some relationship with Gardner's categorisation of reasons for stylistic judgments about music, and also to some extent with Copland's three planes of Listening, but they seem even closer to the four phases of aesthetic development in Art, Drama and Music proposed by Ross (1984).

Ross labels his four phases as Displacement, Improvisation, Convention and Composition, and although his explanation is based upon the similarity of practical activities taking place in each of the Arts areas at the same ages, it is also implied that aesthetic appreciation follows the developmental structure of complexity set by these phases.

For Music, Displacement is the phase, from 0-2 years, when sensory sound explorations set up the memory patterns which enable recognition and discrimination to take place. Improvisation, from 3-7 years, occurs when tunes are recognised as such, and are developed through doodling and repetitive practice. Musical memory continues to develop leading to the ability to anticipate the next phrase of the music. Also in this phase, there is found a delight in pure sound and its use for sound effects. This phase could be held to parallel Parsons' stages of Beauty and Realism, and Expressiveness. Ross' next phase, Convention, which also seems to mirror Parsons' Style and Form stage, occurs from 8-13 years, and this is where a desire to conform to the adult musical world emerges and manifests itself in the desire to play a musical instrument, to play and compose in familiar idioms, and to listen to popular music. Autonomy emerges in the final phase, Composition, which Ross speculates as beginning at about age 14. In this phase students develop a sense of music as being a *"form of communication and language of personal expression"* (Ross 1984 p.130) and come to accept that a piece of music is a symbolic structure. It is at this phase that the search for quality begins and a taste for the more complex forms of music emerges. *"All the previously acquired skills of discrimination, of encoding and decoding, now come into their own in the pursuit of the*

transcendental, the visionary, the personally expressive." (Ross 1984 p.130) This last phase clearly mirrors Parsons' stage of Autonomy, which stresses the search for quality and the importance of re-examining and questioning the accepted classifications and judgments in order to form a personal set of values.

The close parallels between these theories, which have been noted here, are not surprising as Ross' theory is based upon observations of student activity in all three areas of the Arts. Parsons' hypothesis, based upon over 300 interviews conducted systematically over a ten year period with subjects ranging from Preschoolers to Art professionals, proposes stages which are basically similar, but which could also be used to evaluate students' responses to visual stimuli in a hierarchical manner. Both these theories also have similarities with a view of the development of musical experiences as a continuous spiral, put forward by Swanwick and Tillman (1986), which seems to expand Copland's three planes into a complete musical theory.

Supported by data from a four year study of the compositions of 48 children aged 3-11, the spiral shows the continuous nature of musical development from the earliest years to adulthood. The spiral consists of eight developmental modes : Sensory, Manipulative, Personal Expressiveness, Vernacular, Speculative, Idiomatic, Symbolic and Systematic.

According to Swanwick and Tillman, the Sensory mode which seems to be prevalent up to 3 years, is characterised by a fascination with timbre and dynamics, and is an era of sound exploration. The next stage, the Manipulative mode is shown in the desire to control sounds from instruments and other sound sources, and it is here that the sense of pulse and beat also emerge. Songs may show the first signs of the development of Personal Expressiveness and it is in this mode that the awareness of phrasing first appears, and speed and dynamic changes are often used to indicate emotions. Established musical conventions become important in the next mode, the Vernacular, which can appear from age 5 and is clearly established by 7 or 8. It is in this mode that cultural conventions become

apparent in the repetition of phrases and patterns derived from the students' aural experiences. The Speculative mode, which seems to develop between 9 and 11, accepts the patterns found in music and attempts to introduce fresh ideas or surprises. Students working in this mode are willing to experiment, trying to integrate their ideas into a recognised structure. By the age of 13 or 14 students are striving to work within a recognised Idiomatic format, and placing their speculative surprises in accepted places, such as the ends of phrases, as they endeavour to conform to a recognised musical style. Meta-cognitive processes emerge in the symbolic and systematic modes appearing at ages 15 and over. At the Symbolic level strong commitments to music, the awareness of its power and the ability to articulate about it emerge, together with the capacity to reflect upon experiences. The "fully fledged musical person" operates at the Systematic level, conscious of stylistic principles, able to consider the possibility of altering them, and able to discuss this rationally and philosophically.

This Spiral Theory relates modes loosely to age levels, and examination of data presented from the four year study reveals that students can be working in up to four different modes at the same age, although the majority of compositions will be in one mode. The available published work (Swanwick and Tillman 1986, Swanwick 1988a, Tillman 1989) details research with students from 3-11 but does not indicate extensive research with students outside this age range, and although mention is made of visits to secondary schools and of the examination of writings by composers, it is unclear upon how much evidence the conclusions drawn with regard to the symbolic and systematic levels are based.

A later analysis of musical experiences by Swanwick (1988b) reduces the levels to five and seems to have even closer links with research into Music Listening. The first two levels here are Attention to the Sound Source and Impression of Sound Materials, which can be related to Copland's sensuous plane, Myers' intra-subjective (1927) and Yingling's sensory classifications (1962) and the vast amount of work on perception of the separate elements of music. The next level, Perception of Expressive Character clearly

relates to the expressive plane of Copland, Myers' associative and character classifications and Yingling's associative and emotional categories. Swanwick's highest levels Structural Framing of the Work and Value Position are operations in Copland's sheerly musical plane, Myers' objective and Yingling's intellectual categories. This response analysis could be used to evaluate student responses if the response fell clearly into one of these categories, and if it is accepted that musical elements can be classified into an exclusive hierarchy like this one.

Swanwick has himself seen the implications of the spiral theory and its derivatives for student evaluation and has published some suggestions for GSCE assessment criteria based upon it (Swanwick 1988a), which are quoted in Appendix B. These criteria are very interesting, but imply that listening ability follows the same continuum observed in compositions, without the overlapping of modes which seemed to be evident in his original research. Each grade relates directly to one of the modes in the spiral, and contains only musical elements deemed to be dealt with in that mode. There seems to be an assumption that performance at a higher level implies competence of all the lower level criteria, but pieces of music have different dominating musical elements, and a student response which does not show awareness of this would be a lower level response than one that dealt with that element in detail. It is the composer who decides the respective importance of each musical element in a composition, and the acute listener will observe and report this. In addition no provision has been made for differences in performance within each mode, so as they stand these criteria might be difficult to implement in the classroom. In all fairness to Professor Swanwick, it must be noted that his assessment criteria were published merely as examples of how a theory could be used as the basis for assessment.

Theories of Cognitive Development

Other researchers, with a psychological rather than musical background, have put forward theories to account for cognitive development. It could be that one of these theories has a hierarchical structure of increasing cognitive complexity which could be used as the foundation for an evaluation technique.

One of the first, and most famous researchers, to draw observations in this area into a coherent theory was Piaget. Piaget's theory of intellectual development (as described by Ginsburg and Opper 1969) puts forward four main stages from birth to adolescence : Sensorimotor (0-2), Preoperational (2-7), Concrete Operational (7-11) and Formal Operational (11+).

In the Sensorimotor stage the child initially uses its reflex actions to explore its immediate world, then begins to control its movements as it recognises visual cues, and by the end of this stage it has developed some memory and can remember solutions to previous problems. The Preoperational stage is the one in which language emerges and memory develops leading to greater control over self and the awareness of social behaviours. Perception at this stage is limited with attention being focussed on the dominant feature of an item rather than considering all aspects. The next stages are those in which school age children are found. In the Concrete Operational stage children characteristically develop the abilities to decentralise their thinking, to consider several items or aspects of an object or problem at once, and to maintain ideas about a substance such as volume regardless of cosmetic changes. Problems can be solved, but success is largely dependent upon the presence of stimuli which can be manipulated. The ability to use abstract, hypothetical reasoning is the main characteristic of the highest stage the Formal Operational stage.

Recently doubt has been cast upon the conclusions drawn by Piaget and the accuracy of some of his testing methods. Alternative tests striving to avoid a verbal bias seem to show that he may have set his age limits too high (Donaldson 1978), and that variation in

performance within specific domains may exist (Lamborn and Fischer 1988). Studies do however seem to bear out the essential principles that cognitive differences exist and that they can be classified into stages.

As the secondary age group is mostly contained within the Formal Operational stage, it might be possible to utilize a theory derived from Piagetian ideas, which further subdivides this stage as the basis for an evaluation technique within this age and cognitive level.

One of these new theories, Fischer's Theory of Cognitive Development; The Control and Construction of Hierarchies of Skills (1980), although not based upon the Piagetian model, consists of a continuum of 10 levels of increasing complexity classified into three broad tiers : Sensory-motor, Representational and Abstract, which can be related to Piaget's stages. The highest level of each tier also forms the lowest level of the next tier, so that the transition between the skills in each tier can be seen to be a smooth natural progression. Unfortunately only two levels, 7 and 8, fall within the high school age range, so despite the fact that these span the transition from Representational to Abstract, the theory does not offer sufficient scope for immediate conversion to an evaluation method that would suit the purposes of this study, though it does indicate that both types of thought might be found amongst students of that age, that abstract thought is the higher of the two, and that there is scope for further research in this area.

More relevant to this age group and perhaps of more use as the basis for evaluation in the high school is the theory of Case (1980, 1985), which has four main stages from ages 0-18 with four substages of increasing complexity within each stage. Sensorimotor is again the first stage lasting from 0-18 months, being a basic development of awareness of the immediate world, in which the child starts to differentiate between objects and actions. The next stage from $1\frac{1}{2}$-5 years is called Relational and is typified by the child's usage of the relationships between objects and actions which were observed in the previous stage. Dimensional, the first school age

stage, lasts from 5-11 and is characterised by the ability to work out from aspects of a problem the significant dimensions of that problem. The stage relevant to high school students is the Vectorial, which deals with the adolescent's developing ability to process simultaneously diverse information from more than one dimension in order to produce abstract thought structures.

Each of these stages has four substages : operational consolidation, unifocal coordination, bifocal coordination and elaborate coordination, which indicate increasing levels of complexity in the thought processes required to solve problems. Case puts forward interesting ideas about the contribution of working memory and short-term storage space towards movement through these substages (Case 1980), and these aspects of his theory are perhaps of particular value for teachers. Although he states that the total structural capacity of the memory remains constant once it is fully developed by about age 2, the ratio between the two functions of the working memory, for storage of problem solving strategies and for controlling the actions involved in the strategy, alters as the child progresses through the substages. In the early stages the greatest amount of capacity is required to coordinate control of the strategy, but as with practice the strategy becomes automatic, less memory is required for attitudinal control and more becomes available to consider additional factors and alternative strategies, and so the child is able to progress into more complex problems and devise more intricate strategies.

In each main stage (Case 1985), operations from the previous main stage are first consolidated and tranferred from attitudinal control to storage capacity, in the Operational Consolidation substage, before moving to the next substage, Unifocal Coordination, where the first aspect of a problem in this new stage can be considered. Once this first aspect has been mastered, then that strategy too enters storage leaving space for the consideration of another aspect together with the now automatic first one, and this is the Bifocal Coordination substage. Repeated experiences at this substage once again lead to reduction of the need for conscious attitudinal control, releasing functional memory capacity for

consideration of multiple aspects and the attainment of the highest substage, Elaborate Coordination. At this substage practice consolidates all the strategies of the main stage until sufficient working memory is available to consider the problems of the next stage.

Like Fischer, Case regards the highest substage of each main stage as serving a dual function. It is not only the highest level of attainment of the earlier stage, but also functions as a transitional plateau of consolidation before further growth continues along the cognitive continuum.

This theory seems to have possibilities for the secondary school situation due to the inclusion of hierarchical substages in the Vectorial stage. It also seems to be more relevant to the classroom situation as, although research projects have shown that movement between stages in different content domains occurs at about the same time, they have also shown that it may occur at slightly different times between content areas and individuals. Case is also prepared to acknowledge that movement between stages and substages can be influenced by maturational and non-cognitive factors such as : motivation, previous experiences, environment, peer and family interactions, and the child's own interests and desires. The child's desire to imitate, to explore new ideas and situations, and to be willing to attempt fresh problems, are also admitted as major affective factors. The inclusions of these environmental and attitudinal factors should make this theory acceptable to teachers, who regularly encounter these influences on children's progress. Despite these positive aspects and although the theory has obvious implications for the organisation of curriculum items and teaching strategies, Case's major published work on this theory (1985) does not present these ideas in a way which is easy to assimilate or adapt to classroom practice. Whereas it might be possible to use this theory in the construction of a technique to teach a specific musical skill, it is not immediately clear how this theory could be adapted to a general evaluation technique for Music Listening, which is the aim of this study.

A theoretical work which is widely used by teachers in training, and therefore known and already adapted for practical use by practising teachers, is the Taxonomy of Educational Objectives (Bloom et al 1956), commonly known as Bloom's Taxonomy. A taxonomy is a classification which shows the natural relationship between items, such as plants, animals or educational objectives. It could be that such a taxonomy could be of assistance in the evaluation of Music Listening by classifying and stating the relationship between types or degrees of responses to listening.

Taxonomies

Bloom's Taxonomy of Educational Objectives (1956) was developed to assist communication between educators, and to facilitate the development of curricula and the evaluation of their objectives. Initiated by a group of American college examiners, it was widely discussed by teachers, administrators and researchers at every stage and revisions were made reflecting the views of as many educational practitioners as possible. It could therefore be said to represent the practices current at the time, and to have been published with the approval of contemporary educators. Initially designed to be a classification of the student behaviours which might represent the desired outcomes of educational objectives, it was not intended to be a value system. Evaluation and grading within each category were to be left to the teacher, though as it eventuated each category did represent a greater degree of involvement with the material, and so a value system did emerge. The wide nature of student activities in the curriculum meant that three broad types of desired outcomes could be identified, and these were the three domains of cognitive, affective and pyschomotor activity.

The Cognitive Domain (Bloom, Engelhart, Furst, Hill and Krathwohl 1956) deals with activities such as handling theoretical knowledge and applying it to solve problems and be creative. It has

six categories : Knowledge, Comprehension, Application, Analysis, Synthesis and Evaluation. Each category has several levels of definition, each of which can be broken down still further into sub-categories.

The Affective Domain (Krathwohl, Bloom and Masia 1964) may seem to be less objective than the cognitive domain as it deals with the student's emotional growth through the development of interests, desires, attitudes and appreciation. All of these may be influenced by cognitive activities, yet the five categories in this domain can be seen to be separate from the cognitive domain. The categories are Receiving, Responding, Valuing, Organisation and Characterisation by a Value or Value Complex.

The Psychomotor Domain (Harrow 1972) deals with the kinetic or movement based development of the student, and although this appears at first to be linked to physical education, it is in fact an integral part of any subject involving the development of practical skills. Technological subjects such as Woodwork, Cookery and Typing and the practical Arts subjects such as Pottery, Drama and Music all involve the learning of many kinetic skills, and they can be introduced into other subjects to cater for children whose prefered mode of learning is through this domain. This domain has six categories : Reflex Movements, Basic-Fundamental Movements, Perceptual Abilities, Physical Abilities, Skilled Movements and Non-Discursive Movements.

The hierarchical relationship between categories in each domain can be seen most clearly in the psychomotor domain, where the controlled locomotion movements of the Basic-Fundamental category are clearly operations at a higher level than the first category of innate Reflex movements such as the baby's tonic neck or grasping reflexes. The aesthetic creative movements of the ballet dancer (Non-Discursive movement) are also the result of more intensive and conscious control and therefore an outcome at a higher level than the ability to catch a ball (Skilled Movements). It is this hierarchical aspect of the taxonomy that suggests it may be more useful for evaluation purposes than an analysis of types of

response.

Each of the three domains is relevant to Music as they are integral to all musical activities. Performance skills on an instrument or vocally take place in the psychomotor domain, though they are governed by knowledge, comprehension and application of fingerings, notation and conventions of repertoire from the cognitive domain and progress is dependent upon affective components such as the value placed upon the activity, which determine the amount and quality of private practice sessions. Composition though essentially a high level cognitive activity involving continual judgments and reanalysis as existing knowledge is synthesised into a new form, also involves skilled practical performance of each trial and the finished product. The affective domain is also involved as all cognitive judgments made during the course of composition are dependent on the values and preferences held by the composer. Listening is also dependent on the affective domain as without a willingness to attend to the music, only passive listening (or hearing) can take place, and active listening is the goal of music education. Once the student has begun to listen, if the piece is recognised then values and preferences may either stimulate closer attention or remove the state of willingness to listen. The affective domain is therefore important to listening, and affective goals are usually included in syllabus objectives for this area. Physical responses probably at the reflex level also exist to listening (cited in Chapter 2, Harvey 1986), however the majority of syllabus aims and teachers's efforts are directed towards the cognitive aspects of listening, and it seems reasonable to expect that the cognitive taxonomy could be applied to music listening.

In the cognitive domain the categories are arranged on a continuum from simple to complex, although the order of categories does not necessararily reflect the levels of cognitive activity involved in each. The subcategories also are numbered, though they are often of equal value. It is therefore difficult to estimate the relative values of outcomes in each category, though there would seem to be a consensus that the first category is the most basic and simplest one.

Knowledge centres upon memory, and can easily be related to music listening as the recognition and identification of sounds is the basis of any discussion of a listening experience. Specific instrumental sounds, conventions and categories shown in styles and periods of music, and the principles involved in performance practice, composition and the science of acoustics are all aspects of aural recognition. Questions set at this level for Music Listening would usually ask *What?* has been heard.

Included in the Comprehension category are translation, interpretation and extrapolation, all of which are concerned with understanding the relationship between aspects identified in the first skill category. The relationship between written notation and music, between the various aspects of a piece of music and the use of this knowledge to lead to expectations about what comes next in the music, such as that a slow movement will follow the opening movement in a concerto and contain a sustained tune accompanied by subdued orchestra, or that the recapitulation of a sonata form movement will seem to mirror the exposition, all come under this classification. These skills may initially seem harder to assess, but they can still be observed in response to questions based on the question *What?*. This type of question could include notation based problems such as *Identify the second theme by marking it on the score*, *What are the major features of the form of this piece?* and *What will happen in the next section of the music?*

In Music, Application is continually being tested, as there are very few musical situations where factual knowledge is required without application to a situation. In Listening, for instance, the identification of a particular artist takes into account many specific facts recalled from previous listening experiences and relates them to aspects of the present performance.

Analysis, breaking down the music into parts in order to discover the relationship between them and the principles upon which it is constructed is an activity commonly found at the higher levels of music education. Questions such as *How?* and *Why?* can be used

both to identify the constituent parts of the music and to establish their relationship to each other and overall organisational principles.

Synthesis and Evaluation are the categories for which the ability to act at all the previous levels is required, and these are perhaps the levels upon which music critics and professional musicians operate. Jazz musicians continually synthesize aspects of previously heard performances into fresh improvised versions of standard tunes, and writers about music incorporate their ideas and those of others into their analyses and criticisms. Evaluation is also present in the work of these people, though it is not always articulated verbally.

Although its original intention was to facilitate discussion of the assessment process, this taxonomy has often been used to set curriculum objectives and to plan classroom teaching strategies, and it is seldom mentioned by teachers when devising assessment procedures. It is a complex classification system, and although it has possibilities for the evaluation of Music Listening, to implement its use for this purpose a comprehensive battery of tests would have to be designed. The use of such a battery of tests in the classroom would occupy an excessive amount of time and could possibly cause the loss of balance between the three areas of music education.

Another taxonomy classifying only the Listening aspects of Music has been devised by Goolsby (1984) as part of an attempt to promote Music Education as aesthetic education. Analysing the separate skills required into subskills of concept to be learned, this musically detailed taxonomy also provides assistance for teachers by listing the difficulties which are anticipated to arise in the teaching of each concept.

Goolsby's first three skills deal with the perception and description, or analysis, of what is heard. The first skill, Perceiving and Describing the Components of an Auditory Aesthetic Object deals with the recognition and identification of the

component elements of sound, such as pitch, volume, durations and sound sources (timbre), whilst the second skill, Perceiving and Describing the Relations among the Complexes of an Auditory Aesthetic Object (regional qualities) is concerned with the relationships between these elements, which he denotes at being the "first-level complexes of music". These complexes are composed of combinations of the basic elements and include aspects such as melody, tempo, harmonic rhythm, tonality and texture. The formal structures found in music are given a separate skill category, Perceiving and Describing the Organization of the Complexes of an Auditory Aesthetic Object. This category has two skills subcategories; one dealing with the interpretation of types of music such as song, which may have extra musical meanings; and the other dealing with *"human regional qualities"* in music, which seem to be the intrinsic 'emotional' responses to sections within pieces, such as introductions, bridge passages and codas. The final skill is Assessing the Aesthetic Value of a Musical Work of Art. Bloom places this skill firmly in the cognitive domain as evaluation, but Goolsby has a wider view of this skill and incorporates both objective and affective reasons into the justifications for decisions at this level.

Responses to music are always influenced by past experiences and the emotional reactions they recall. It could be that this occurs because music is experienced at a primeval, physical level as well as an intellectual one, so the use of the cognitive domain of Bloom's Taxonomy could be too limiting. Goolsby's taxonomy, whilst based in the cognitive domain, and admitting the inclusion of the affective domain in the evaluation of Music Listening, seems, like Swanwick's spiral to be basing the evaluation of music upon a hierarchical view of the elements of music. A typical student response to a piece of unfamiliar music such as *"That's no good because I don't like it"*, where the student is expressing their idea of the work's value without seeming to be aware of the basic elements or their relationships, is not a high level response despite being a value judgment and therefore in Goolsby's highest skill category. Perhaps both Goolsby and Swanwick's classifications might be valuable for use at the tertiary level, where all responses

might be expected to be objectively based upon musical evidence, but they don't seem to be immediately applicable to the qualitative evaluation of school students' responses to music.

Both the taxonomies considered, Bloom and Goolsby, could however be effective if used in the planning of classroom activities, though they seem to be cumbersome, perhaps too complex, and inappropriate for use as Music Listening evaluation tools. A third taxonomy, the SOLO Taxonomy (Biggs and Collis 1982) stemming from the same premise of the increasing complexity of thought, and based upon extensive observations of classroom responses in many subject areas, seems to have evolved a simpler overall response classification system which is perhaps more relevant to the classroom situation.

The SOLO Taxonomy

The SOLO Taxonomy is named for its content. the Structure of the Observed Learning Outcomes, and originated in studies which examined secondary school student responses in several subject areas with a view to revealing the Piagetian developmental stages which they displayed. As it became clear that response levels in individual students were not stable, but could fluctuate according to the subject matter or situation, the focus in the study moved from developmental stages to learning outcomes. It seemed that perhaps the hypothetical cognitive structure of an individual and the actual responses they gave were not necessarily the same thing, and that previous experiences and learning might account for the differences which were being found.

The taxonomy was developed to focus upon the structure of the response presented by the student to a set task, and to provide a mechanism to evaluate the quality of learning by examination of this structure. Quality of learning is revealed in the hierarchical classification of five types of thinking, which range from

Prestructural, the lowest level where the student does not even attempt to complete the task, to Extended Abstract, the most complex level, which is the recognition of new, flexible, original thinking. The levels in this theory bear a close relationship with Piaget's stages, but the SOLO levels are not intended to provide a stage label for a student, but to describe that individual's performance on a specific task at one particular time. Levels are not held to remain stable within an individual across different content areas or across tasks within an area, as experiences relevant to each specific task will vary influencing the depth of thought and therefore the SOLO level.

Figure 3.01 presents a metaphor which shows the basic principles of this taxonomy in diagrammatic form. The stimulus data is shown in the the centre of each diagram as solid circles bounded by horizontal lines, with extra data which is relevant but was not given in the stimulus represented by hollow circles below, and irrelevant data shown as crosses above. Lines connect the subject (S) to the data which has been noted, show the relationships noted between the data items, and link these to the subject's response (R) on the right hand side. These lines clearly show the increasing complexity of the response structures.

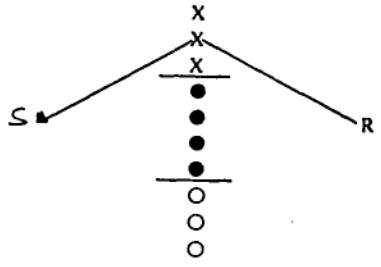
As can be seen on the diagram, the Prestructural level is characterised by a response which does not mention any of the data given in the stimulus. The response uses a data item not given, which is inappropriate to the task set. This type of response does not show that the subject understood the question, and may indicate either inability or unwillingness to complete the task set.

Unistructural level responses show evidence of having understood the task but use only one aspect of the data to complete it. They ignore other features given and so may therefore jump to conclusions, which would be refuted by other features given in the data.

Many aspects of the data provided as the stimulus are noted at the Multistructural level, but they are presented as separate items

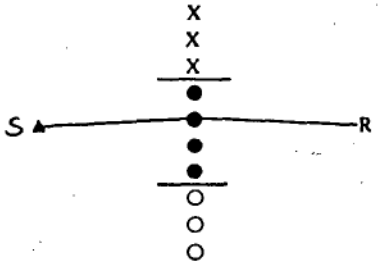
Figure 3.01 The Principles of the SOLO Taxonomy

(from Biggs and Collis 1982)



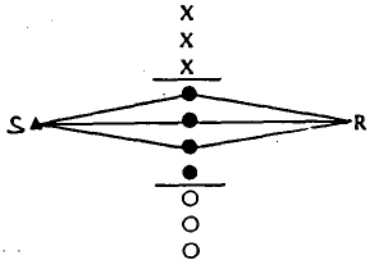
Prestructural

Doesn't understand question
May restate question
Guesses - no logical basis



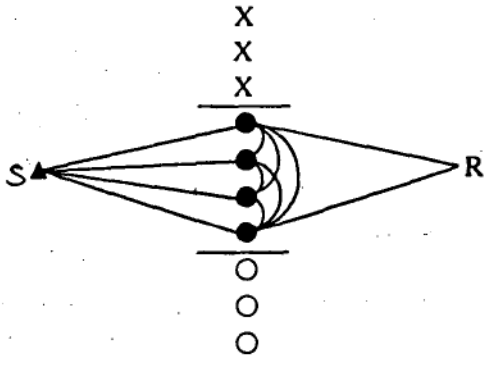
Unistructural

Answer based on one feature
Conclusions may be incorrect



Multistructural

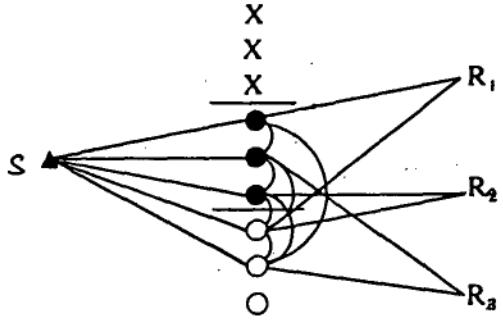
Several features mentioned but not linked together
Isolated series of judgments each based on one type of data
Same material may give conflicting conclusions
Quantity increase from Unistructural



Relational

Overall concept or principle forward to account for the data
Data elements related to each other and to other materials previously learned

Solution sticks within the known



Extended Abstract

True logical deductive process
Extra evidence produced
Abstract principles tested and used to make predictions
Full conclusions not needed
Original thinking, not limited by bounds of present knowledge

Answer shows intellectual courage and confidence

and possible connections between them are not noted. As a characteristic of this level is the inability to retain large amounts of facts simultaneously in the memory and relate them to each other, it may seem as if conflicting ideas are being presented, and the response could seem like a series or list of separate answers.

Responses which link ideas together are at the Relational level, and it can be seen on the diagram that lines connect separate items within the data in different ways, showing their various relationships to each other. At this level competent responses are provided, which make generalisations based on accurate observations and interpretations of the given data. Deductions are made which are true within this context, but which might not hold true if extra data were to be considered.

The highest level of responses, Extended Abstract, has the most complex structure. Here, data additional to that provided in the stimulus but relevant to the task set is included in the response. As in the previous level relationships between the existing data are noted and fresh data is tied into these relationships. The student working at this level demonstrates the ability to generalise and make hypotheses outside the range of ideas presented, does not feel the need to keep within strict limits or to make definite decisions. Responses also demonstrate the student's willingness to be open to ideas, to explore ideas and accept the possibilty of alternative solutions, and they may present several alternative solutions.

This hierarchical structure has parallels with Bloom' Taxonomy in that the levels begin with simple thought directed at one idea and proceed to more complex thinking involving a synthesis and evaluation of many ideas and basic principles. It is also similar to Case's concept of the increase in complexity through substages in that each level shows an awareness of more aspects of the problem. This could, as Case asserts, be due to the changing structure of the working memory. SOLO is however a much simpler system, which has in addition already been adapted to the evaluation of a multitude of school subjects, including some like Poetry and Creative Writing,

which are traditionally thought of as being difficult to assess fairly. It seems reasonable to expect therefore that it could also be useful in the evaluation of Music Listening responses.

If SOLO can be used to evaluate Music Listening responses then it should be possible to find responses to Music Listening experiences which are suitable for each level. To obtain some preliminary data to ascertain whether this could be possible, the investigator collected the following responses to music from secondary school and tertiary students and classified them into SOLO levels using the general principles stated above.

The lowest SOLO level is the Prestructural one, where the given response is inappropriate to the question. These responses to a Music Listening task could be classified as Prestructural :
"Nothing, I can't hear anything."
"I don't like this music, why don't you play Bon Jovi?"
"Talking about music is too hard."
Each of these responses ignores the intent of the task and does not try to answer it.

Where the response mentions only one aspect of the problem it is classified as Unistructural. A genuine, but limited, attempt is made to answer the question, and the response may be quite elaborate. Unistructural Music Listening responses might mention only one aspect or element of the music, and be similar to these collected responses :
"There are lots of instruments playing. Clarinet, violin, flute, trumpet, drums and trombones. It has bells in it. The violin plays a lot. I like lots of instruments playing."
The only aspect mentioned in this response is instrumentation.
"This is a loud piece of music. It has lots of loud bits in it. It starts with a loud bit. There's one soft bit and lots of loud bits."
Volume is the sole aspect considered here.
"This piece is a scene on a river bank, with people feeding ducks, just like at Richmond in the summer. There are ducks on the grass asking for bread and some ducks swimming fast up the river, hoping to get some before it runs out. The music makes me see this picture

in the music."

No musical elements are used in this description, only the images conjured up by its expressive aspects.

Each of these four responses could be a response to the same piece of music. They may be honest attempts to answer the question, but because each mentions only one aspect of the music, none of them gives a balanced or clear description of the piece, and each should be classified as a Unistructural response.

Multistructural responses on the other hand mention several features of the stimulus, so a Music Listening response would identify many aspects of the music, and therefore give a clearer picture of it. As a characteristic of this level is the inability to retain many facts simultaneously in the memory and relate them to each other, each aspect noted would be mentioned in isolation, and it might also seem as if conflicting ideas are being presented. A Multistructural response could seem like a series or list of separate responses.

"The music makes me think of a field where rabbits are playing together. They are all very happy. It is happy music. The sun is shining."

This response uses two elements contained in the music, the imagery of the rabbits in the field, and personal mood.

"The music is fast and loud. It is happy, dancing music. The violins play. Violins, flutes, trumpets and piano and a high instrument. It is a sad song."

This is also a multistructural response as many musical elements are listed : speed, volume, mood, imagery, instrumentation, pitch. The apparent discrepancy between "happy" and "sad" exists because whilst both descriptors are true, they occur at different times within the music and no explanation has been given that the music contains separate sections within it.

"The music is fast and played by lots of instruments. It has some good tunes. I like the violin tune at the start. It is happy music that dances along. It reminds me of a scene in an olden time ballroom with ladies in long dresses twirling around. The middle section seemed slower. There was a loud bit. The trombones had a tune. It seemed to be a happy piece."

This multistructural response mentions speed, instrument, several tunes, mood, imagery, volume and also has some insinuations as to period and formal structure, but as no connections between them are made, it is classified as a Multistructural response.

Higher Relational level responses, should show and explain an understanding of the relationships between different aspects of the stimulus in a logical way with all statements supported by evidence from the music. These responses show these characteristics.

"This piece sounds just like music that they would use to try and scare people, perhaps in a film. It sounds scarey as its all disjointed, no instrument ever plays a long tune or blends with others. The notes keep changing from high to low as well, and soft to loud, which stops you from adjusting and getting used to the music, and so keeps you scared."

Here, several aspects instrumentation, pitch and volume have been drawn together to explain one central idea, the mood created by the piece.

"This piece of music is extremely expressive and it brings to mind vivid pictures of death, disaster loneliness, The drums sound like guns in the distance. The harmony made by individual notes blending for a short while and then suddenly clashing into dischords and the patterns which the instruments use add to the effect and feeling of evil."

Once again the imagery created by the music has been explained by describing several aspects of it : mood, instrumentation, texture, harmony, melody.

"The piece began with a pleasant sweeping flute tune accompanied by the mainly string orchestra. This opening was in $^{3}_{4}$ time with a lazy rhythm that made it sound like an olden times waltz from a film where the ladies wear long dresses that swirl around as they dance. Suddenly there is an interruption as the brass come in with loud heavily accented notes which upsets the people dancing, this is the middle section of the music. This bit is loud and fierce with the main tunes played by deep brass instruments, probably trombones, euphoniums and tubas. It builds up to the fierce section gradually and so also seems quite menacing. Happily it doesn't last long and the peaceful dancing continues as the violins take up the main tune

again."

This whole response centres around an awareness of the form of the extract and is a description of the main elements within each section. Imagery is used to clarify the style of the opening section, and the reasons for this choice of image are clearly given.

Extended Abstract is the highest level, at which the material provided as the stimulus is synthesized together with other knowledge and basic principles to devise fresh possibilities. A musical response at this level should use the musical stimulus as a stepping stone for the evolution of basic principles that could refer to other pieces or areas of music. Responses at this level could lead into areas not required by the question and so might seem to contain irrelevancies, but these should be set up as a direct result of exploring the question and evidence thoroughly. The following response has clearly gone beyond the bounds of a Relational level response.

"This Romantic orchestral extract reveals the same melodic, formal and orchestration techniques used by the English composer Elgar in his symphonic writing, and could possibly be an extract from one of his orchestral suites. As it has strong dancelike characteristics, it is probably a dance movement.

Elgar, like artists in the other art forms in the 19th and early 20th Centuries was very much influenced by the Romantic movements stress of the beauty of nature, and he, like other composers elsewhere, such as Sibelius in Finland, liked to get close to nature by frequent long outdoor walks. It seems that the freedom of nature is often expressed in their music. The melodies of this piece, played predominantly by strings doubled with woodwind in the standard romantic manner, swoop and soar like the birds Elgar watched in his walks over the Malvern hills. Legato phrasing is used to create this smooth flying effect and it also makes the less conventional intervals which he used seem totally acceptable, singable and normal. Unlike Romantic composers like Berlioz, who wrote long yearning melodies, Elgar often has short phrases within his melodic structures as in this extract and this extensive use of legato ....."

This response describes the music in an almost cursory manner by

identifying its chief characteristics and placing them in their context. It then goes on to isolate one aspect, the type of melody, and explain this by suppositions as to the personal habits of the late romantics in general and their influence on their music. The second paragraph appears at first to be irrelevant, but it is linked closely to the music in the third sentence, and the response is unfinished, perhaps indicating that the writer felt more could be said on the topic.

Each of the five SOLO levels reveals a new stage in comprehension and thinking, and these sample responses show that they can be applied to Music Listening. In drawing parallels between SOLO and Piagetian based developmental stages, Biggs and Collis relate the SOLO Multistructural level to the Middle Concrete stage at ages 10-12, the Relational level to Concrete Generalisation at 13-15 years and the Extended Abstract level to the Formal Operational stage expected after 16 years. This would seem to indicate that these three SOLO levels might be encountered in the responses of secondary school students. When intending to use SOLO with this age range therefore, tasks should be designed to stimulate these levels and aim towards the higher levels, Relational and Extended Abstract. Tasks must be therefore carefully structured, so that instant one word solutions (Unistructural level response) are recognised as being inadequate, and students understand that the response requires not only recall of previous knowledge, but also comparison, discrimination, recognition, clarification, classification, review and restructuring of that knowledge. This analysis technique could be seen as handling the incorporation of cognitive domain requirements, whilst leaving room for elements from the affective domain, so it seems to be suited for the kind of open and speculative questions with which the Arts deal.

SOLO and Music

Are there any connections between SOLO and Music? Could there perhaps be a natural relationship between them? Some of the literature was re-examined to determine if there were any relationships between SOLO and the purely musical response analysis systems, which might support the proposal of using SOLO for the evaluation of Music Listening.

Swanwick's spiral, referred to earlier in this chapter, does not at first glance appear to bear any similarities to SOLO, but as the first two modes Sensory and Manipulative seem to be concerned with the successful achievement of one aspect of music, they could perhaps be loosely regarded as being Unistructural in nature. Depending on the number of elements used to create emotional effects, the Personal Expressiveness mode, where emotions are represented by changes in elements, could be considered as either Unistructural or Multistructural. Awareness of musical styles, which are created by the many musical elements in specific ways, could also be regarded as Multistructural, whilst the Speculative mode, in which students are working towards the effective placement of new ideas could be either Multistructural or Relational. The Idiomatic mode is the one in which new ideas are integrated into established formats, so this must show evidence of the Relational level of thought. As Swanwick claims that the Symbolic and Systematic modes are the ones where personal reflections and the discussion of possibilities of altering established principles occur, these seem to be the modes which would display Extended Abstract thought.

Composition, the highest phase presented in Ross' theory of Aesthetic Development (1984), also seems to have similarities with the Extended Abstract level, and Parsons' theory of stages in the development of Art Appreciation (1987), also shows definite changes in thought towards the Relational and Extended Abstract levels in its later stages, Style and Form, and Autonomy. There are therefore possibilities that both these theories could be discussed in terms of SOLO levels.

A study using methodology similar to that of Parsons' was that of Gardner (1973). The comments on the reasons given for stylistic recognition in this study could also be classified into SOLO levels. The 6 year olds were either unable to give a reason, which is a Prestructural level response, or used one aspect of the music, a Unistructural level response, whereas the 8 and 11 year olds who mentioned several features of the music were responding at the Multistructural level. Different strategies were employed by the adolescents and college students tested, who pointed out the relationships between the various musical elements which gave the pieces continuity. As Gardner put it they *"had a sense of a musical piece as a structured entity which possesses continuities, direction, possibilities and implications"*, which would seem to indicate thought at least at the Relational, and possibly at the Extended Abstract levels.

The responses printed in the account of Nelson's work (1984) with violin students also seem to fall naturally into the SOLO categories, Unistructural and Multistructural. Responses to the questions *Do you like this song?* and *Why?* elicited these responses : *"Yes, because I like it"* and *"Yes, it has a lot of E string in it"*, which, as they mention one aspect of the music only, are clearly at the Unistructural level; and *"Yes, it has a nice rhythm and beat"*, which is a two element Multistructural response.

Goolsby's Music Listening Taxonomy (1984) can also be seen in terms of SOLO levels. His first skill, recognition of the components of music, could be either a Unistructural or Multistructural level response depending on the number of components identified, whilst the next two skills concerning the relationships and organisation of musical elements are clearly at the Relational level. His highest skill, assessing the aesthetic value of the piece of music, would probably be at the Extended Abstract level, as it would require comparisons to be made between the musical stimulus and other pieces of music, musical styles and interpretations.

The classification of aural test items according to degress of difficulty carried out by Colwell (1987, cited in chapter 2) also seems to show parallels with SOLO. The easier tasks, such as identifying instruments and counting phrases, involve only one musical element and could therefore be regarded as requiring the Unistructural level of thought. Being aware of the progress of many musical lines is necessary for the correct identification of texture, so this harder test item could be placed at the Multistructural level. The next level of difficulty included matching sound with notation, which seems to be a Relational level skill as it requires the integration of many musical concepts into a complex symbol system, which then has to be simultaneously related to a transient patterns of sounds.

There does seem to be some evidence here that not only could SOLO be used to assess Music Listening, but that through the parallels found between the work of music theorists and researchers, this usage could possibly be justified on musical grounds as well as psychological and practical ones. It was therefore the application of the SOLO Taxonomy to the evaluation of Music Listening in the classroom which this study set out to investigate.

Chapter Four

The Pilot Study and the Development of the Test

The Pilot Study had two purposes, to determine whether or not student responses to pieces of recorded music could be analysed using the Solo Taxonomy and, if this were possible, to develop a technique for the classroom assessment of listening to music using the principles behind the SOLO Taxonomy. This classified student reponses as Prestructural, Unistructural, Multistructural, Relational and Extended Abstract. A one year Pilot Study was used to establish that student responses could be analysed using the SOLO Taxonomy, and develop an assessment technique, which was later used with 328 students in Grades 7-10 over a three year testing period. In the assessment technique students listened to three extracts from recorded music and gave written responses to an open-ended question about the music. This chapter will explain how this technique was developed through the Pilot Study, and then used in the classroom with the Pilot Study group.

It was anticipated that a number of questions would be resolved, or partially resolved by the end of the Pilot Study, which would allow procedures for the Main Study to be finalised. These questions were :

1) Can student responses to music be assessed using SOLO ?
2) Is it possible to ask questions about music that stimulate responses at all the SOLO levels ?
3) What listening assessment procedures can be used with SOLO ?

It was anticipated that a number of questions would be resolved, or partially resolved by the end of the Pilot Study, which would allow procedures for the Main Study to be finalised. These questions were :

1) Can student responses to music be assessed using SOLO ?
2) Is it possible to ask questions about music that stimulate responses at all the SOLO levels ?
3) What listening assessment procedures can be used with SOLO ?
4) Is there a listening assessment procedure suitable for use with a class that can be used with SOLO ?

It was also hoped that indications towards solutions of the following questions might emerge :

5) Is it possible to use any piece of music as the stimulus?
6) Is it possible to analyse written responses efficiently using SOLO ?
7) Are student performances influenced by factors other than Listening ability?
8) Is a SOLO level based assessment fair to the students ? Is it comparable to other listening assessments for these students ?

During the Pilot Study classroom assessment procedures for listening were observed, and two were selected for further investigation. They were developed using SOLO and trialled with students from Grades 7, 8 and 10. The observed responses and reactions of these students were critical factors in the final selection of the testing technique, which was trialled with a group of Grade 10 students. Following discussions with music teachers and this final group of students, some modifications were made to the test paper and its administration before the Main Study commenced.

Figure 4.01

SOLO Stages Related to Music in Responses to Music

SOLO	Music
Prestructural	
Doesn't understand the question May restate the quesion Guesses - no logical basis for response	Didn't listen to the music May latch onto the obvious in an attempt to prove they were listening but without relating it to the question
Not a serious attempt to answer the question	
Unistructural	
Answer based on one feature Conclusions may be incorrect	Only one musical concept used Answer is relevant to the question but limited in scope
An attempt has been made to answer the question	
Multistructural	
Several features mentioned but not linked together Isolated series of judgments each based on one type of data A quantity increase from Unistructural Same material may give conflicting conclusions	Several musical concepts used but not linked It could be a series of separate answers Quantity of detail may be impressive Abstract concepts may be mentioned but not explained in the context of the music
An honest attempt to give a good answer	
Relational	
An overall concept or principle put forward to account for the data Elements of the data combined and related to each other and to materials previously learned Solution sticks within the known	More complex musical concepts used which link aspects of the music and show their relationship to the question A competent workmanlike solution to the question which shows good musical knowledge and awareness
Conclusions may not be true outside this situation	
A competent answer in which points made are proved by evidence from the data	
Extended Abstract	
True logical deductive process Extra evidence is produced Abstract principles are presented, tested and used to make predictions Several hypotheses proposed and tested Full conclusions not necessary Response may be characterised by "maybe, probably, possibly, it may be that.."	All elements of music used to stimulate thought Thought is extended beyond present piece of music Answer goes further than required An open-ended answer - leaves room for possible contradictions and opposing opinions Musical elements that may be used as the basis of response : performance practice, interpretation, relationship between stylistic elements structural complexities and debates.
Original and speculative thinking, not limited by the present bounds of knowledge	
This answer may be considered irrelevant unless the question is sufficiently open	

Theoretical Relationship between SOLO and Music Listening

The first step taken towards the development of the test was to develop the theoretical relationship between SOLO stages and Music further. This theoretical relationship was developed along the lines outlined in Chapter 3. The musical side of this model, which evolved from the study of student responses to Music Listening and the application of these to the general SOLO response model, is shown on Figure 4.01.

At the Prestructural level a response is characterised by not mentioning any of the data given in the stimulus, so for Music Listening a Prestructural level response was taken to be one which did not provide any evidence that the student had listened to the musical stimulus and related it to the task set. Obvious elements or aspects of the music may have been mentioned or insinuated, but if there was no connection to the question asked, then the response would be assessed as being at the Prestructural level. The implementation of this definition required that the wording of the question should clearly show that the answer should make reference to elements within the music.

As Unistructural level responses show evidence of having understood the task, but use only one aspect of the data to complete it, if one aspect of the music was mentioned and related to the intent of the question, then the response was deemed to be at the Unistructural level. The essential difference between the Prestructural and Unistructural levels here, was that at the Unistructural level, the student had attempted to link their aural observations to the task set, whereas at the Prestructural level there had been no such attempt. A Prestructural response could appear to be independent of the question, whereas a Unistructural response would be dependent upon it.

Many aspects of the data provided as the stimulus are noted at the Multistructural level, but they are presented as separate items and possible connections between them are not noted. So for a Multistructural response classification, a student should have observed and presented several musical concepts as solutions to the set problem. These might be presented as a series or list of separate responses. The musical concepts mentioned could be ones that implied a view of the musical stimulus which linked concepts to each other such as formal structure, but if their relationship to the stimulus and each other were not fully explained, and the series of items appeared to be an unconnected list, then the responses would be classified as Multistructural.

Responses which link ideas together to solve the set problem are classified as being at the Relational level, and so a fully Relational response could focus upon an abstract musical concept such as formal structure and style to provide a clear picture of the musical stimulus, and offer some informed suggestions as to its background. Relational level responses would be those which demonstrated that the students had observed the music in some detail from a broad perspective, and were thus in positions from which they could make informed judgments about it. The presentation of the response should demonstrate the student's background knowledge and an understanding of the relationship between it and the present musical stimulus.

The highest level of responses, Extended Abstract, has the most complex structure with data additional to that provided in the stimulus, but relevant to the task set, being cited to create and support hypotheses. An Extended Abstract response therefore would require the incorporation of data from the musical stimulus into an overall picture, which might create the possibility that a Relational response could be confused with an Extended Abstract one, as both would require the demonstration of knowledge of other pieces and styles of music and extensive general musical knowledge. However as the essence of an Extended Abstract response is the production of thoughts, in this case about the music, which were original to the student, an Extended Abstract response would speculate about the

original stimulus and deliberately cite evidence from other pieces of music, styles, composers etc. in support of a reasoned argument. It would not be restricted by the original stimulus, but would extend the bounds of the original question, expand into general theories about music in general, and perhaps postulate some questions of its own. It might not present a definite solution, but present several alternatives and open up avenues for further exploration. Aspects of music which could possibly arouse this type of debate were considered, and the following were considered: performance practice, artistic and technical interpretation, the relationship of stylistic elements, and aspects of composition such as formal structure, texture, orchestration and harmony.

Figure 4.01 also indicates that the increasing structural complexity of the response through the levels, may involve not only higher level cognitive operations upon the data, but also musical concepts of an increasingly complex and abstract nature. This is particularly noticeable at the Relational level where the more complex musical concepts which link aspects of the music may form the basis of the reponse, and at the Extended Abstract level where abstract principles form the basis of speculations, and it is interesting to note that this to some extent mirrors the hierarchy of musical elements *("components and complexes")* put forward by Goolsby (1984), and described in the previous chapter.

The musical side of this model, which seemed to have a relationship to Goolsby's skill levels, was then refined into guidelines for the classification of Music Listening responses into SOLO levels. These are shown on Figure 4.02, and were used throughout the Pilot and Main Studies as the basis for making decisions about SOLO levels.

Figure 4.02.

Guidelines for SOLO Assessment of Music Responses

Prestructural

An irrelevant answer.
No musical elements mentioned.
Doesn't seem to understand the question.

Unistructural

One musical element is mentioned.
It may be dealt with in some detail.

Multistructural

Many aspects of the music are mentioned.
It may be a competent answer, but it is presented as an unconnected list of items.

Relational

A clear workmanlike answer, demonstrating understanding of musical concepts and their relationship to each other.
Statements are proved by direct reference to the music.
Response is presented as a logical connected sequence of ideas.

Extended Abstract

Evidence of in-depth thinking putting forward original ideas and incorporating musical knowledge other than that within the musical stimulus.

Selection of the Test Procedure

Once the evaluation guidelines had been established, the next step was to find a classroom assessment procedure for which they could be used effectively. Listening, an essential component of all musical activities, is present in all musical outcomes, and can be assessed formally through written tests and assignments, and informally through teacher observations of student contributions to class discussions, solo and group performances and creative activities such as improvisation, arranging and composing. In the practical situation it is almost impossible for the observing teacher to determine how much of the student's response is due to their own listening awareness, and how much is due to their ability to copy the ways they hear sounds produced by others, so for the purpose of this study a small aspect of listening awareness, which could clearly be attributed to one individual, was separated from other musical activities and examined. This, the student's response or commentary on music they had just heard, was also the aspect of Listening commonly isolated by teachers for assessment purposes at Moderation meetings.

All classroom activities that involved students actively listening to music were examined for their possible use in an assessment technique. These included both oral and written procedures. The criteria by which these procedures were examined included factors common to all effective testing situations, some peculiar to the music listening experience, and suitability of the procedure for analysis of responses into all five SOLO levels. To ensure that a test can be used to compare individuals, it must contain the same material, be presented under the same conditions without external interference, and be scored in precisely the same way for every student. For an aural stimulus the quality of reproduction must be high and it must also be equally audible to all students, so that the same venue, equipment and seating arrangements are commonly used for audio testing. External distractions must also be eliminated as far as possible, as they divert attention from the stimulus and prevent the music being retained in the memory for

later comparisons and comment.

Test anxiety is another factor that can seriously affect the performance of some students causing them to give responses which are not representative of their normal achievement level. Test anxiety should be reduced as far as possible so that students give responses more closely related to their normal levels. A problem commonly found in music tests, which causes student anxiety is that the pace of a Music Listening test is determined by the speed of the music and cannot be controlled by the student or test administrator. The pace of the music is always too fast to allow students to write as they listen, so to reduce anxiety and allow the music to be effectively absorbed into the memory, the test should include a time for undisturbed listening and the music should be repeated. As the purpose of the test was to become a tool for classroom music teachers, perhaps one of the most important criteria for selection of a procedure was the practicality of its administration to a class of 20-30 students.

All these criteria were considered when examining the assessment procedures currently in use in the classroom for their suitability for use with SOLO. The oral procedures considered included individual interviews, whole class and small group discussions, and student talks about music that they had heard. Written procedures considered included assignments, written worksheets, concert reviews, live performance criticisms, and formal tests requiring one word responses, multiple choice answers, free response and essay type answers. Those procedures that were thought to be suitable were trialled using the principles of the SOLO Taxonomy. (Details of these investigations are in Appendix C.)

After considering all the procedures available and trialling superitem worksheets and free responses, it was decided that a free response to an openended question which could be answered at any level, would perhaps be the most suitable method of assessment. More practical for testing large numbers of students than an interview, it would give each student the freedom to select when to write, when to listen for clarification, to select their viewpoint, to expand

and clarify their ideas. Though an interview might get reactions that were closer to the listening experience, a written response would be less likely to interfere with internal recall of the music and would give the student time to consider their response. It would also allow each student to respond at the level they thought appropriate for the question and the piece of music, and not try to force them to answer questions at all levels as the superitem worksheets had done, thus causing distress and motivation problems for later tasks with some students. Other advantages of this method included the fact that it was similar to normal classroom procedures and could quite easily be done under test conditions so that the responses were definitely the unaided work of each student. Although it would maximize as far as possible the number of students to be tested, there were some disadvantages. These included the possible strong influence of linguistic and writing skills and the inability to challenge or clarify statements.

Trials of the free response procedure were undertaken with a class of 15 Grade 10 students. This class included two outstanding students both musically and academically and two low ability students, one of whom was semi-literate, so the class could be said to be representative of the wide spectrum of abilities normally encountered in music classes of this age group.

As 33% of final marks were determined by their performance in the Listening section of the course, these students normally listened to a wide variety of music, discussed it and wrote about it either in class under supervision or for homework, so that it was possible to use the tasks and questions normally set as part of the trials. Some of the usual questions offered scope for many response levels, though many did not seem to provide encouragement for Extended Abstract responses. Thus some tasks needed rewording to provide adequate scope and encouragement for the higher levels of Relational and Extended Abstract thought. It was found that endeavouring to include scope for these higher levels whilst not making the tasks too daunting for lower ability students was quite difficult.

The Question of the Question

Questions which endeavoured to determine which aspects of the music the student had focussed upon by demanding descriptions seemed to be suitable for use with SOLO, and when trialled, these did produce responses from Prestructural to Relational. Some students, whose responses were fully at the Relational level, put forward original ideas based upon their knowledge of other music, which seemed to be possible starting points for Extended Abstract thought. The responses were not however sufficiently developed to be considered true Extended Abstract responses. The questions which stimulated these responses mainly dealt with abstract music, and included these :

"Describe how the opening tune alters in mood each time it reappears by explaining the changes in the orchestration, texture and piano part."

"This music has clearly defined sections within it . Describe the melodic structure, texture and orchestration of the main sections and compare this style of composing with that of other composers."

"Describe 2 sections from this piece."

"Describe this piece of music as clearly and fully as possible. Compare it to other music you have heard or played. You may like to comment on these aspects of the music :

instruments	*patterns of tunes*	*types of tunes*	*speed*
rhythm	*volume*	*harmony*	*mood*
style	*texture*	*orchestration "*	

"Describe the main episodes in this piece showing how the instruments and rhythms chosen by the composer alter the mood."

Some questions, like the last one above, demanded a response which required the linkage of several musical ideas into one concept, and would therefore be at the Relational level. Yet despite this, these questions were answered by some students at lower levels, Unistructural and Multistructural. Even questions in two parts, like those following, which define the task clearly and sometimes have extra information given, were not always answered at the Relational level.

"This piece is like a conversation between 3 voices. Describe each voice clearly and by reference to the structure of the melodies explain the type of thoughts each voice is expressing."

"This piece is typical of music written at a certain time in history. Which period was it ? Explain what it is about the music that made you come to this decision and compare it to music of other periods."

There was also a tendency amongst a few students to respond to the music by creating stories or describing the mood of the music and the images it created in their mind, rather than describing the musical aspects of the music as required by the question. Whether music can or should communicate feelings and images is a contentious issue, and attitudes towards it vary in different cultures and groups in society. However it is apparently a common reaction amongst Western students as Bula and Szymanowski reported in 1987 that even absolute music such as Chopin's Mazurka Op 59 No 3 was in most cases *"often interpreted by young people in a literary or picturesque mood with a strong emotional implication"*. This investigator found that this expressive and associative reaction to music also came from some adults with minimal musical training, who were amongst the group of adults asked to listen to the final test music. Questions were therefore devised which acknowledged the likelihood of imagery being perceived and guided students firmly towards a more musically based response which would show the relationship between the images and the musical elements of the music :

"What feeling or emotion does the music suggest to you ? Which musical elements create this expressive effect ?"

"This piece of music is meant to create the picture of a person. Listen carefully and describe what sort of person you think the music depicts. How has the composer used the music to create this image ?"

"The composer is trying to create a picture in sound. What do you think his picture is ? Why do you think that the music suggests this picture ?"

"..... wrote ..... to create the impression of .....

How has he used the instruments and music to do this ?

Do you think he has been successful ?
Could the piece be improved ? How ?"

These questions were successful in leading most of the students, whose previous response to music had been merely to tell a story, into an attitude where they focussed their responses on the musical elements which had created the mood, character and imagery evoked for them by the piece. It seemed that some pieces, especially those that might be thought of as programmatic, initially stimulated more emotional and associative responses than others, and that it was necessary to provide an outlet for these student reactions in order to get them to consider the music from a purely musical point of view.

It was also noted that the questions *How?* and *Why?* stimulated responses at the levels P, U, M and R, but that recall type questions such as *What?, When?* and *Who?* tended to produce more one word Unistructural responses. These lower level questions did however motivate the lower ability students to make more effort and incorporating this type of question into a larger task encouraged these students to make more effort and try to produce longer responses. Questions structured in this way did not present the task as a large global problem but as a series of small manageable steps:

"What kind of scene or mood does this music evoke ? How has the composer created this mood ?"

"This piece of music was written for a specific purpose. What do you think this purpose was ? Explain how the different aspects of the music led you to your conclusion."

These two part questions were particularly useful with programmatic music, but could be applied to all types of music and as can be seen from the sample questions above, they clearly demand responses at the Relational level. Students however often gave Multistructural responses, stating their reasons, but not connecting them to the question by explaining their role in the formation of their judgments. It was not clear on paper whether they had omitted this because they felt it was too obvious or because they had not made the connection. When questioned about this, some students gave

verbal answers which showed that their written statements were the basis for their judgments but that they had not realised it was important to connect the two, and some gave answers showing they had not realised the full implications of the question. It seemed that there could be some ambiguity in Multistructural responses between the intentions of the student and what they wrote down.

Types of Music used for the Stimulus

Most of the music used with the Grade 10s was written between 1830 and 1985, as the Romantic and Modern periods of musical composition were the focus for study at this point in the curriculum. This meant that an extremely wide variety of music ranging from Tchaikovsky through Bartok to Schoenberg, Sculthorpe, jazz, country, and pop music could be used. Some pieces written before these periods were also included as part of their work on stylistic recognition.

The trials conducted with these students showed up a number of trends about their responses to different types of music. Students found pieces in familiar idioms easiest to talk about as they had already developed a suitable vocabulary, and had learned to listen for the most important aspects of that style. This particular group gave good responses to Baroque music, which they had studied in the previous year and although they initially found Romantic orchestral music hard to discuss, they soon learned how to handle it competently in terms of its orchestration. Even after intensive tuition and practical performance and composition experience, however, they continued to experience great difficulty with atonal and electronic music, which were styles completely new to them.

Vocal music led to responses which focussed exclusively upon the meaning of the words and ignored all musical elements of the composition. The students seemed to have difficulty listening to anything other than the lyrics, and even when a song in a foreign

language was presented they tried to work out "what the song was about" from the words rather than using the highly expressive accompaniment. Vocal music was therefore ruled out for use in the main study, because of the students' inability to focus on the musical rather than literary aspects of this type of music.

Pop music also stimulated poor responses as most students had deeply held preferences and beliefs as to which styles and specific performers were acceptable. They were unwilling to listen in depth to their favourites, prefering to sing along rather than listen, and they were reluctant to hear other pieces unless they were presented in a historical and sociological context as with "Rock Around The Clock" and the development of the rock music culture, and so pop music was therefore also not considered for inclusion in the test.

Very poor responses were also stimulated by the various types of Non-Western music (Aboriginal, Japanese, African and Indian) to which they had been introduced during the year. Some students had difficulty even accepting that these unfamiliar idioms could be classified as music, and found it extremely difficult to be sensible when talking about them. It seemed as if longer responses, which would give each student a better chance to attain a representative level, would be gained by using more familiar Western music in the Main Study.

During the Pilot Study students were often asked, after they had made their responses, whether or not they had liked the pieces, and it was noticed that students often made the least effort and produced their poorest responses to pieces they disliked. Even if their writing skills were very limited they made more effort and produced better responses for pieces which they they liked. The least liked pieces did however produce the widest spread of responses, as the lower ability students gave less effort and the higher ability students continued trying to answer the question objectively. This fact was pointed out to the students during their lessons, and they were continually reminded to try and ignore their personal feelings and be objective about what they had actually heard.

Development of the Procedure for Analysis of Responses

During the Pilot Study responses were assessed in several ways: SOLO categories alone; an Elements Checklist; a numerical scoring system; and revised SOLO categories, which indicated Transitional responses. Experienced classroom music teachers were also involved in discussions at each stage and eventually a final method emerged which used a combination of aspects of the trials, and was acceptable to them all. It was this method that was used in the Pilot Study final test and the Main Study. Initially responses were allocated to a SOLO category using the guidelines presented earlier in this chapter. All responses were also checked for mistakes and given an assessment using the current Schools Board of Tasmania award system, details of which are in Appendix A. This was done so that the students could be given a form of feedback which would provide information on their progress, whilst not influencing future responses by alerting them to the criteria behind the SOLO assessment.

It was soon noticed that responses which demonstrated the more complex levels of thought, and were therefore classified as Relational, also typically seemed to refer to more complex and abstract musical elements, such as formal structure, whilst the simpler elements, such as speed and dynamics commonly seemed to form the basis of responses classified at the Unistructural or Multistructural levels. It seemed possible that the theoretical basis of increasing complexity of both cognitive operations and subject based concepts, outlined in Figure 4.01, was being borne out in the practical situation. Responses to a piece of music used earlier were re-examined and this observation was confirmed as the musical elements used did become more complex as the operational level of thinking about the music became more complex. A decision was therefore made to examine the elements of music and see if they could also be classified using SOLO.

In consultation with 4 secondary music teachers a comprehensive list of elements that could be mentioned by listeners of all abilities was drawn up and is shown on Figure 4.03. With the exception of the last group of elements, technical criticism of performers, the completed list was similar to those compiled for use by music researchers such as Hevner (1956), Getz (1966), Flowers (1983), Hair (1987) and writers on music such as Anthony Hopkins (1979). Many texts for music appreciation and active listening have also been written based on these elements, such as those by Colles (1919), Thomson (1978), Bamberger and Brofsky (1979) and Sadie and Latham (1985).

This list was then sorted into three main groups as shown on Figure 4.04. The first group contains elements that can be identified by themselves, and the second contains musical elements whose recognition is dependent upon the identification of several elements from the first group. Goolsby (1984, cited in Chapter 3) refers to these second group elements as "complexes" and "structures". For example, identification of a specific instrument is a basic element, but awareness of that instrument's usage in combination with others to create a specific sound is orchestration, which is a higher level concept as it requires the listener to combine and evaluate not only instrument recognition but also the effects of timbre. A third group was also found to be necessary, and into this were put elements such as interpretation which would be mentioned only by those who had a firm understanding of the second level elements and their interrelationships, and it was this group of elements which it was expected could lead to original thoughts. These three groups of musical elements have a clear affinity with the SOLO levels - Unistructural, Relational and Extended Abstract.

Figure 4.03

List of Musical Elements that could be Mentioned in Responses to Music (drawn up in consultation with music teachers)

1: Instrument or Voice Identification

soloists group genre (type of piece)

2: Orchestration

timbre/tone colours blends of sounds

3: Mood Indicators

speed/tempo dynamics/volume metre

rhythm accents syncopation

4: Melody

melodic shape - rising/falling - smooth/jagged - scalic/chordal

phrasing

articulation -legato/staccato

5: Form

melodic repetition sequences climaxes ostinato

contrasts

musical forms sections within the music

6: Texture

monody homophony polyphony/counterpoint elements

accompaniment use of silence

7: Tonality

mode/scale used harmony/chord progressions

stylistic background

8: Style

historical/idiomatic style of performance(performance practice)

9: Technical Criticism of Performers

intonation tonal quality control of instrument

use of contrasts stylistic elements tempo excitement

relationship between performer/rapport director

Figure 4.04

A Proposed Hierarchy of Elements mentioned when Listening to Music

Group One

Instrument and Voice Identification
Group of Performers
Speed
Volume
Melodic elements : phrasing,
  shape of tune, articulation, accents
Rhythm and beat
Mood and Imagery

These are elements most likely to be found forming the basis of Unistructural and Multistructural responses.

Group Two

Orchestration
  timbre/tone colours, blends of sounds
  use of sectional groupings
Formal Structures
  genre (type of work), musical forms,
  melodic repetition, sections,
  sequences, climaxes
Texture
  accompaniment, monody/homophony/polyphony,
  counerpoint, use of silence
Tonality and Harmony
Style
  historical, idiomatic

These are more complex musical structures which would possibly form the basis of Relational level responses.

Group Three

Criticism / Interpretation elements
  stylistic elements (performance practice)
  selection and use of contrasts
  tempi, dynamics, intonation,
  tonal quality, control of instrument,
  creation of sense of excitement
  rapport between performers
  direction
Principles of Composition

These elements could form the basis of Extended Abstract responses.

The first group of elements on Figure 4.04 included elements which required only aural recall skills for recognition, and if these were mentioned alone, the response would clearly be classified as Unistructural. For instance, recognition of speed requires identification of the frequency of the beats, and recognition of an instrument requires identification of the sound quality, thus only one aspect of music is needed for recognition of these elements. Identification of the quality of some elements may require the comparision of the present sound to ones previously heard, as in the case of volume, but this is still a thought process concerned with the recognition of an isolated musical element. No elements were found to require separate recognition of two aspects without any recognition of their relationship to each other being needed for identification, so there were no Multistructural classifications.

For inclusion in the second group of musical elements an element must require the collation of aspects of two or more elements for recognition to occur. For the recognition of tonality or formal structure previous passages of music must be recollected in order for comparisons with previously heard pieces of music to be made. Combining observations of several less complex elements is a Relational level skill. This also happens in the recognition of styles such as Dixieland Jazz. The recognition of this style for instance, requires the differentiation of specific types of several musical elements such as orchestration, formal structure, harmony, rhythm, melodic pattern, accent, articulation and instrumental technique, and an understanding of their interaction before the style can be accurately identified. Recognition of a musical style is therefore clearly a Relational level skill.

On Figure 4.04 timbre, the recognition and identification of the quality of individual sounds, is listed in Group Two despite being a Group One element. It was placed in this group because of its intrinsic relationship to orchestration, a more complex element. The contributory features of formal structure such as melodic repetition were also placed in the second group although they are not themselves complex musical elements. This was done to simplify the use of the checklist.

Two aspects of listening considered as being likely to lead towards responses dealing with universal and basic principles of music, requiring the incorporation of external knowledge, and comparison with other listening experiences which have possibilities for extension far enough be classified as Extended Abstract, were performance criticism (including performance practice) and the principles of composition. These two aspects of discussion about Music were therefore placed in the third group of elements.

This SOLO classification of elements does not necessarily indicate the SOLO level of the whole response as elements may be handled in different ways. Elements from the first group are prerequisites for any discussion of Music, and could quite logically be introduced as part of a Relational or Extended Abstract response. Multistructural responses could be built from a combination of elements from each group, which if presented without logical connections could not be considered as a Relational response. This classification of elements does not therefore provide a method of SOLO assessment, but as it gives an indication of the levels of thought involved in the recognition of various musical elements, it could be of assistance in the initial recognition of levels of thought and depth of musical perception involved in a response to a piece of music.

This proposed hierarchy of elements (Figure 4.04) was circulated to the same group of 4 teachers for judging on a purely musical basis. At this stage they knew nothing of the SOLO basis for the groups. They were asked whether the elements had been sorted into three groups of equivalent degree of difficulty (with no ranking within the groups) and if the hierarchy between the groups was one they would use when marking student work. They were also asked to make any additions and alterations that they thought necessary. All the teachers consulted accepted this hierarchical grouping of musical elements without feeling the need to make any changes, most thought it would be very helpful when marking student work, and some asked to keep their copies of it so that they could use it for this purpose.

This hierarchy was then set out as an Elements Checklist for student responses. At the first reading through of a response each element was noted as it was mentioned. When allocating a SOLO level it was noticed that Relational responses always mentioned at least one element from Group Two, whereas Unistructural responses were usually confined to Group One elements. Multistructural responses might draw on elements from both groups, but tended not to discuss elements from Group Two in terms that proved the musical elements were understood. Thus the use of this Checklist confirmed the previous observation about the increasing complexity of both the level of thought and the musical elements used in the responses.

This Elements Checklist (Appendix F) was also used as the basis for an experimental scoring of responses. As elements in each group could be used at any level, scores for each element were weighted from 1-4 according to the way the element was used in the response. Unistructural and Multistructural usage scored 1 point, Relational 2 and Extended Abstract 4. Without the weighting this was successful as a measure of variety in a response, but even with the weighted scores it was unfortunately not accurate as a measure of quality. It was possible for Multistructural responses containing extensive lists to gain a higher score than better quality Relational responses which were clearly thought out and well argued, but dealt only with a few pertinent features of the music. This marking scheme was not only a inadequate measure of quality but also cumbersome, and time consuming to use, although it gave greater freedom and scope for original thought than current essay marking schemes.

Some responses were also found which, whilst clearly belonging in one of the SOLO categories, were better than other responses in that category. At the Unistructural level these responses dealt with the musical idea in greater detail and sometimes at great length. The better Multistructural level responses either presented many more points than other responses, or seemed to be verging on the brink of the Relational level as they introduced and explained the importance of a more complex Group Two musical concept, such as orchestration, but then reverted to listing other points. There were

also Relational responses which introduced an idea outside the music and made a brief speculative point about it, but did not then develop this idea fully. These responses seemed to be in a state of transition between stages, and although they were classified into their main level, they were also given a subsidiary classification indicating quality within the level.

During this initial phase of the Pilot Study short experiments with different assessment procedures were carried out, which culminated in the selection of the free written response as the test medium. These experiments are described in Appendix C. Continual observation of students' attitudes and reactions towards the tasks set had also been carried out to ascertain which tasks would stimulate the best responses. All written responses had also been analysed, not only to determine if students were having any problems with the instructions given, but also to determine whether the use of SOLO as an assessment procedure was appropriate in these situations. Enough evidence had been gathered to start constructing the test to be trialled with the Grade 10 Pilot Study group and then used in the Main Study.

Construction of the Test Format

All decisions made about the test format were based upon previous music research findings and observations made during the Pilot Study. A written free response format was selected because it gave scope for responses at all levels, was simple to administer, was a format familiar to students and could be used as a normal end of unit listening assessment test. To reduce student fatigue as far as possible, facilitate test administration and make it as near a normal test format as possible, the test length was set at 40-45 minutes, the approximate length of a school period.

In a normal Music Listening lesson students usually only dealt with one piece of music at a time, and it had been noticed that if a student did not like the piece presented they did not put much

effort into their responses, omitting many aspects of the music and working at a lower level than usual. As, when questioned, they revealed that they had in fact heard these aspects quite clearly and understood that they should have mentioned them in order to answer the question fully, it seemed clear that their dislike for the piece had caused their poor written response. It was therefore decided to provide several pieces in the final test so that student preferences and prejudices would not affect their overall score. There would also then be the possibility of disregarding the worst score when trying to estimate the student's level. It was also felt that if the pieces were representative of a variety of styles, then students would have the opportunity of finding at least one piece for which they would want to give a good response. Research using actual pieces or extracts from pieces of music had used between 1-25 extracts (Flowers 1983, Gardner 1973), but the number of pieces had always decreased sharply with the amount of writing and type of judgments required from the subjects. For descriptive writing 1-3 extracts (Valovy 1981, Bula and Szymanowski 1987) seemed to be the normal number.

Research by Bartlett (1973) and Bradley (1972) had shown that repetition was an important factor in the development of liking for a piece of music, so it was decided to repeat each piece three times during the test. This would also allow each student time to think, to verify and extend their original observations, and time to construct their response. If the test were to be conducted within a 40-45 minute time allocation, which would have to include adequate time for explanation of the test procedure, playing the extracts and student response writing, each extract would have to be quite short. It was decided that three extracts each lasting about two minutes would fit into the time allocation.

Selection of the Music for the Test

The selection of the three pieces of music to be used in the final test was a complex and lengthy process involving an investigation of results obtained by other researchers from various types of music, and many hours of listening to select extracts of the right type and length.

Getz (1966) had found it better to use works unfamiliar to the students so that they could be led to an intellectual consideration of the work rather than repeating associative meanings, learned reactions or prejudices (Bula and Szymanowski 1987). Short pieces of music tend to be familiar to students as they are often used for signature tunes and by teachers as creative stimuli, and they also tend to be programmatic. It was therefore resolved to use extracts from larger works, which would each have to offer enough variety to allow students to comment on many aspects and produce Multistructural responses, yet have some obvious main concepts such as formal structures, compositional principles and stylistic elements relating to performance practice which could be grasped by the more aware and lead to Relational and possibly Extended Abstract responses. Getz (1966) and Gardner (1973) both also considered the quality and recording standard of the music used to be important, so only works by composers accepted as master craftsmen, and high quality recordings were considered for inclusion in the study.

When students in the Pilot Study had been presented with several pieces, a tendency was observed for some of the students to describe the first piece in depth and then to compare the resources used to perform the other pieces to the first one rather than discuss any other elements in the following pieces. It was resolved to avoid this problem by ensuring that all three extracts were performed by the same musical medium, and in this way it was hoped that students would focus their attention on each piece as an individual piece rather than pointing out the obvious differences between them.

As students had also been observed to pay more attention and to give better quality responses when they liked the music, the research on musical preferences was examined. This revealed that orchestral music was the likeliest type to be acceptable. Zumbrunn (1972) had noticed that adolescent boys paid more attention to orchestral music with woodwind instruments in it, than to strings alone. As it was also reported that music played by instruments the listeners had themselves played was also well received, it was decided to use orchestral music featuring wind and brass instruments, which these students had all experienced playing, as the stimulus. Factors given for liking music in research by Getz (1966) and Prince (1972) included fast, lively tempo, repeated diatonic melodies, changes in volume, and flowing or driving rhythms. Studies by Steck and Machotka (1975), Radocy (1982) and Herberger (1987) also reported preferences for intensity, novelty and complexity, and these were all taken into consideration in the criteria used to select the music used for the test.

As there were to be three pieces it was considered that one of them could contain some of the factors found by these researchers to give negative reactions, such as a jumpy melody, dissonance, atonality and a slow speed. Although the test was not concerned with preferences, and in fact the students were encouraged to be objective in their responses and ignoring their personal opinions, it was felt that it would be interesting to see is a disliked piece would stimulate a different type of response. It had also already been noticed that pieces which were disliked seemed to stimulate a wider spread of responses than pieces which were liked.

Criteria for Selection of Music

Essential Criteria

Orchestral piece featuring wind and brass
High standards of composition, performance, recording
Self contained extract approx 2 minutes
Unfamiliar to students
Identifiable changes in melody, orchestration,
texture, mood, dynamics
Simple form

Additional Criteria for Liking

Fast lively tempo

Diatonic

Consonant harmonies

A wide variety of orchestral music was examined and eventually three passages were found that seemed to fit all the essential criteria and did not seem to have too many different sections within the two minutes : Telemann Concerto in F for Recorder and Bassoon 2nd movement, Varèse Ecuatorial, and Elgar Bavarian Dance No 3 Op 37. Each piece came from a different compositional style and period in the development of the orchestra, and so had obvious differences which could form the starting point for discussion at the Relational or Extended Abstract level. Each had well defined sections of the right length which contained a variety of easily identifiable features to provide key ideas upon which to build responses at any level. The Telemann movement is built upon patterns of swift continuous scalic passages with alternations between the two soloists and orchestra. The conventional tune plus accompaniment texture as found in most jingles predominates the Elgar, with a graceful lilting $^{3}4$ opening rhythm and a loud aggressive brass dominated middle section. The Varèse is a complete contrast containing some of the elements reported to cause dislike. It has no recognisable or singable melodies, dissonant chords, disjointed phrases and individual instrumental sounds instead of conventional orchestral groupings. The musical features of these pieces are detailed in Appendix D.

As the order of pieces could affect the attitude and motivation of the students this was given careful thought. The piece by Elgar, which was similar to the music used in many television commercials and films and would therefore sound familiar and reassuring to students, was placed first. The piece by Varèse, which contained the features noted in the literature on preferences as being least liked, was placed second; and the Telemann, which contained the highest proportion of elements on the list of those liked most, was the clearest to follow, and had fewest obvious changes, was placed third. As it was necessary to include a full repeat of the section

for this extract to make sense musically, this last extract (Telemann) was also the longest. It was thought that because of its extra length the Telemann piece might have to be changed. The piece was however apart from this suitable for the purposes of the test, so it was decided to use it in the Pilot Study test, and observe student reactions to its longer length.

Once the musical stimuli had been decided, attention was then turned to the all important question of the question. From the wide variety of questions used in the Pilot Study, one which gave students freedom to choose whichever musical elements they wished to discuss was selected for use with two of the pieces in the the final test (Elgar and Telemman). This question was

"Describe this piece of music as clearly as possible"

When used with music in idioms similar to the Varèse, this question had not been very successful, as student responses had mostly told stories that they thought would fit with the music rather than describe the music itself. The reason for this could have been that the music was outside their normal experience of music to listen to, but was very similar to music they often heard creating an atmospheric background for films and television programmes. A more definite instruction was therefore devised for this piece of music, which whilst accepting their desire to connect the music with visual images, yet clearly indicated that they were expected to focus their attention on the purely musical details :

"The composer of this music is trying to create a scene or mood in your mind. In one sentence explain the scene or mood the music creates for you. Explain why the music creates this impression for you."

As this question demanded a Relational level answer, there was also an expectation that this question might stimulate more Relational responses than the other questions.

Both the music selected and the questions to be used were discussed with music teachers, and were trialled with two adult non-musicians before they were trialled with the Grade 10 students. Music teachers considered the test to be comparable in type and difficulty to listening tests they had used with students of this

age and experience. The adult nonmusicians, who had received less musical training than the Grade 10 students, but who both had tertiary qualifications, had difficulty responding to the music in any terms other than imaginal ones, and they commented that they felt they lacked an adequate music vocabulary to describe the music. Also, although keen to help with the project, they reported afterwards that they had felt rather threatened by the task set.

The Final Test Format

The test was designed to keep as far as possible within the normal test procedures encountered by students in a school setting. This was so that the findings might be rendered valid for use in other school classrooms.

One of the aims when designing the test paper had been to make it easy to use for students with poor written skills. There were therefore separate test papers for each piece of music, so that students did not have to decide where to start the next question. As little information as possible (Name, Age, Grade, Month and Year) was required to be written at the top and questions and instructions were both written clearly on the paper. The answer sheets were A4 size to encourage longer responses, but as in Yingling's 1962 test, where responses were written on a half sheet of paper, those students who were only able to write a short response were made to feel that this was acceptable as lines were only ruled on the top half of the sheet.

Previous work by Greer, Dorow and Hanser (1973), Greer, Dorow, Wachhaus and White (1973, Greer (1977), and Lamborn and Fischer (1988) had shown that conditions of high approval and support were a positive influence on levels of students' attention, created a favourable attitude towards new types of music, and increased levels of student motivation and performance, bringing each student closer to attaining their optimal performance level. A deliberate attempt was therefore made by the tester to show a positive attitude towards the students and demonstrate high approval of all their queries and responses by smiling and nodding.

During the test the students were seated approximately one metre apart facing towards the loud speakers, thus giving each student freedom from disturbance and the influence of other students, and equal access to the sounds. Each subjects was initially provided with the answer sheet for the first piece only. They were told that they would hear three pieces, each of which would be played three times, and that they should listen carefully so that they could answer the question as fully as possible. They were instructed not to write during the first hearing, but to use this hearing to gain an overall impression of the music. Once this hearing had finished they could jot down points to remember and start writing their answer. The music would be replayed after a short time (equivalent to the length of the extract so that immediate aural recall would not be interfered with) so that they could check any details during the repetitions. During the last writing time for each piece the next answer paper was handed out. All previous answer papers were left on the desks so that rereading and revision could be carried out. Extra time at the end was also given when students were instructed to check their spellings and grammar carefully.

Before each piece was played for the first time, the question, which was also written at the top of the answer sheet, was read out loud, and the students were reminded that anything they wrote would be correct so long as they proved everything they said by reference to the music. Each of the extracts was presented by a number to remove any possible advantage given by familiarity with that composer's music, the period, style, or knowledge of the formal structure revealed by the title e.g Telemann : Concerto, and to remove possible limitations caused by a programmatic title e.g. Varèse : Equatorial.

This final test was trialled with the Grade 10 class at the end of the school year as part of the end-of-course assessment programme. The effect of influences, apart from familiarity with the musical idioms involved, upon responses were also taken into consideration when the test was devised.

Possible Influences on Student Responses

The use of a control instrument was considered, but as the standardised listening tests focus in detail upon the separate elements of music, often use audio signals rather than music, and are therefore not relevant for daily classroom usage, it was decided not to use one for a control test, but instead to examine some school-based factors that might be thought to have an influence upon student responses on this listening test.

As has already been discussed in Chapter 2, there are many diverse factors that may influence a student's response to a listening experience, and it would be impossible to isolate and control for each of these. It was therefore decided to focus upon factors that could be identified, and might to some extent be controllable within the school situation. The school-based factors that it was thought might have some influence upon a written Music Listening response included :

skill in identifying musical elements and talking about music
musical ability
ability to express own thoughts effectively on paper

Each of these factors could have been assessed by tests, but it was considered that these would be intrusive in the atmosphere of the music classroom. To ascertain whether a factor influenced responses, it was felt that an examination of the responses of students identified as being at the extremes in these factors should be sufficient, and that this information could be obtained from observation of the student over a period of time by someone in close contact with them. It was therefore decided to ask teachers to identify students who were either outstandingly strong or outstandingly weak for these factors.

Although teacher judgments of their students have often been shown to be inaccurate when used predictively, they are commonly accepted as the source of assessments of a student's level of achievement at a particular moment in time, and it was on this basis that teachers were asked to identify students. It was noticed that

teachers could identify students at the extremes quite quickly when given their own class list to use, but that they were less able to place students when using a Music group list. They were therefore always given a copy of their own class list and asked to indicate on it only those students whose skills for that factor were either outstandingly strong or outstandingly weak. The names of students who were identified, but were not in the elective Music class were ignored.

Skill in identifying musical elements and talking about music was labelled Listening Ability and was assessed by the classroom music teacher on the basis of spoken responses to direct questions and contribution to classroom discussion in all areas of the Music course. Musical ability is generally taken in Tasmanian High Schools to be the aptitude for and success in playing a musical instrument, so this was labelled Performance Ability. Itinerant instrumental teachers collaborated with the classroom Music teachers on the identification of students for these factors. The ability to express ones own thoughts effectively on paper was labelled Written Fluency, and English teachers were asked to identify students who were outstandingly strong or weak at this skill.

Results of Pilot Study Final Test

During the annual end of year examination period the 15 Grade 10 students who had been involved in the Pilot Study trials were tested using the final test. Response levels to this test were in many cases higher than those observed during the course of the preceding year in similar types of written responses produced by the same students, which may have been due to increased motivation for end of course testing or experience gained during the year. There were no Prestructural responses, Unistructural responses were only found for the second piece and there were equal numbers of Multistructural and Relational responses.

Despite the questions all including lists of musical elements that could be used, responses were often limited in the aspects of the music considered, as this representative sample of responses at

each level shows. (All responses are printed with the original spelling and grammar.) A larger sample of responses for each piece is presented in Appendix E.

Piece No 1 Elgar : Bavarian Dance No 3 Op 27

Question

Describe this piece of music as clearly as possible. You may compare it to other music you have heard or played. You may comment on these aspects of the music : instruments, patterns of tunes, types of tunes, mood, speed and rhythms, harmony and historical background.

"The instruments that played mostly throughout the tune was violins and wind, string orchestra and bass drums and a few percussion instruments. The mood was happy and in the middle of the piece, it played into a clamax as if an important event was taking place. The speed and rythm were medium to fast, the rythms were fairly jumpy."

This response used speed, instrumentation, mood, imagery, form and rhythmic aspects, but presented them as a list, and so the response was classified as Multistructural.

Piece No 2 Varèse : Equatorial

Question

The composer of this music is trying to create a scene or mood in your mind. In one sentence explain the scene or mood the music creates for you. Explain why the music creates this impression for you.

You may mention these aspect of the music : instruments and how they are used, patterns of tunes, speed and rhythms, harmony and type of tunes.

"This piece is boring. It sounds like a slow death march, with soldiers walking into a battle knowing that they're going to lose. Someone is dreaming about it and he was the only survivor and had to go to all their funerals. At the end it sounds as though he just passes away in his sleep."

Unistructural Responses are characterised by the use of only one aspect of the stimulus, and the only musical element mentioned in

this response is speed.

Piece 3 Telemann : Concerto for Recorder and Bassoon

Question

Describe this piece as clearly as possible.

Mention those aspects of the music which strike you as important from this list : instrumentation, formal pattern of tunes, melodic shape, speed, rhythm amd mood, harmony and historical perspective.

Compare it to similar pieces you have heard, and try to work out who wrote it.

"This piece is played by an orchestra with a duet between a flute and a bassoon, these two have interchanging parts which are tangled together. The orchestra begins the piece but after the first short section plays background when these two begin. Occasionally the orchestra again takes over the lead but is put back in its place by the bassoon who again plays in the duet..."

This extract from a Relational response to this piece shows the competent integration of the musical conflict between soloists and concertante with the texture of the part writing.

These extracts from responses to each piece show how the SOLO levels were identified. Within the Multistructural classification there were different types of responses, showing transitional stages within this level, and it could be that this may have been the reason for the larger number of responses classified at this level. It could also be possible that transitional stages within this level are necessary for progression to the next Relational level, in the same way that Case's Theory requires progression through substages within each main stage. Despite transitional responses being noticed, they were not identified separately in the data, but were included in the total for the level in which the response was based, as it was felt that the information provided by the main SOLO categories was sufficient for the purpose of this study.

As a check on the reliability of the definition of the SOLO levels used as guidelines for the classification of student

responses, and as part of an examination of the validity of using this technique to assess Music Listening, 2 experienced secondary music teachers were asked to evaluate some sample responses. Nine responses were used, each being by a different student, and with three responses being selected from each of the pieces used in the test. To avoid any prejudices due to knowledge of the students or the written presentation of the response, the sample was typed and labelled by numbers. The judges were asked first to mark and rank the three responses to each piece using their normal marking system. then the principles behind SOLO were explained, the guidelines for SOLO assessment as stated in Figure 4.02 were produced and discussed, and they were asked to classify each response using the SOLO levels. In every case the allocation of SOLO levels was the same as that of the investigator, and the rankings obtained by the use of the teacher's usual assessment system produced the same rankings as the SOLO method. Both teachers found the SOLO assessment easy to do, and felt that it could be an effective way of assessing student responses.

Data from the Pilot Study

A brief analysis of the data for this test is presented here, and a fuller analysis can be seen in Appendix F. As can be seen from Table 4.01 all students had at least two responses at the same level, either Multistructural or Relational, and four student response sets were all at the same level. Only 3 students produced a Unistructural level response, and in each case this was for the second piece, Varèse.

Table 4.01

Individual Responses Obtained from the Pilot Study Final Test

Students with each Response Patterns

Main Levels					Total
Multistructural	MMM 1	MUM 3	RMM 2	MRM 2	8
Relational	RRR 3	MRR 2	RMR 1	RRM 1	7
					15 students

When the responses were sorted into levels and pieces as on Table 4.02, it could be seen that there were equal numbers of responses classified at the Multistructural or Relational levels. This table also seems to show differences in the pattern of responses to each of the pieces, and to the second piece, Varèse, in particular. There had also been an expectation that the Varèse would produce the greatest number of Relational responses and although it did produce one more response at this level than either of the other pieces, it was also the source of the only lower level, Unistructural responses.

Table 4.02 Classification of Pilot Study Responses (actual numbers)

	Pieces				
Levels	Elgar	Varèse	Telemann	Total	Percentage
Prestructural	0	0	0	0	0
Unistructural	0	3	0	3	6.67
Multistructural	8	4	9	21	46.67
Relational	7	8	6	21	46.67

X^2=37.16 at df=4 significant beyond the 0.001 level.

The nonparametric X^2 Test, used for statistical analysis as the SOLO classifications are only on a loose ordinal scale, revealed the significance of the observed difference to be beyond the 0.001 level of probability. Individual response patterns were then re-examined to reveal the source of the differences. As all students had at least 2 responses at the same level, the response which had the different level was examined. For 7 students the level that was different was lower than the others and for 4 students it was higher. Only one student had a different response level for the Telemann, whereas 4 had different levels for the Elgar, and 6 for the Varèse. Although there were 2 more students with different levels for the Varèse than for the other pieces, this difference was not large enough to enable any conclusions to be drawn. Further work with a much larger sample seemed to be necessary, before any judgments could be made. However as only one student had their different response to the Telemann, which had been the longest extract, and no responses had mentioned its excessive length, it was decided to retain this piece in the test.

Listening Ability

In an attempt to ascertain whether the test was measuring the same skills as the teacher, the responses of six students, who had been identified by their music teachers as having listening skills which were well above or below the average were examined and compared to those of the rest of the class. The comparison of the responses of these students, which seems to show that high SOLO levels were given by those identified as having outstandingly good listening skills, is presented on Table 4.03.

Table 4.03

Individual Response Patterns Classified According to Listening Ability

Listening Ability Groups		
Low	Average	High
MUM	MUM 2	RRR 2
MMM	MRM 2	MRR 1
MRM	RMM 2	
	RMM 1	
	RMR 1	
	RRR 1	

Although the differences between the three groups visible on Table 4.03 were confirmed as being significant beyond the 0.001 level of probability by use of the X^2 Test, and the Low and High group response patterns do seem to be different, in fact there are similarities between each of these groups and the Average group response levels. The majority of the High group responses were at the Relational level, but this level of response was also produced once by a student identified in the Low group for this skill, and there were also 12 Relational level responses produced by students who had not been identified as outstanding for their Listening skills. Although the High group had produced mostly Relational level responses, it seemed that in this group it was possible for students in all Listening Ability groups to achieve high levels. It could be that the trend towards Relational levels, which seemed to emerge

from these results, could be attributed to the continual trialling of listening materials and methods which this group had undertaken during the year. Further work with students less familiar with this type of listening test might give different results.

Written Fluency

As the test was a written response, it was expected that those with better writing and linguistic skills might have an advantage over less fluent writers, and students exceptionally good or bad in these areas had been identified by their English subject teachers. From this group 4 students were identified as being outstandingly above average, and 3 as being well below average in this area, and the response patterns obtained from these students are isolated on Table 4.04 below.

Table 4.04 Individual Response Patterns Classified According to Written Fluency

Written Fluency Groups		
Low	Average	High
MUM 3	MMM 1	RRR 2
	MRM 2	RRM 1
	RMM 2	MRR 1
	RMR 1	
	MRR 1	
	RRR 1	

Once again although there seemed to be differences between the three groups, a closer examination revealed that the same type of response patterns were found in both the Average and High groups. Although the Low group students had produced the only Unistructural responses, all the students in this group had also achieved the Multistructural level for two pieces, as had 5 of the students in the Average group. As the Unistructural responses had all occurred for the same piece, Varèse, it could be that the Low group students had failed to understand the more complex two part question, so that

their lower response levels on this piece might be due to their reading and comprehension skills rather than to Written Fluency alone.

During discussions with the music teachers who taught these students for other areas of the Music course, another possible causal factor emerged, this was motivation towards Music lessons. Each of the 3 Low Written Fluency group students were unwilling to complete tasks set in class, practise their instrument or take part in performances, whereas the High group were conscientious students and enthusiastic band members, who were continually practising and performing. It could be that the three Low group students were not well motivated enough to make the extra effort needed, either to tackle the more complex question, or to overcome their prejudice towards the style of music. It seemed as if there might be several reasons to account for the Unistructural responses noticed in the Low group here, and it was resolved to attempt to resolve some of these questions in the Main Study by identifying students separately for Written Fluency and Motivation towards Music lessons.

Performance Ability

Another factor which it had been expected might influence the quality of responses was Performance Ability. The results of 8 students who had been identified by their instrumental teachers for their Performance Ability were isolated on Table 4.05, and then examined.

Table 4.05 Individual Response Patterns Classified According to Performance Ability

Performance Ability Groups		
Low	Average	High
MUM 2	MUM 1	RRR 2
MMM 1	MRM 1	MRR 1
RMM 1	RMM 1	
MRM 1	RRM 1	
	RMR 1	
	MRR 1	
	RRR 1	

Yet again the levels obtained by each group seemed to show slight differences possibly indicating that there could be a relationship between the two skill areas. The High group had mostly Relational responses, and whilst the Low group contained most of the Unistructural reponses found in this part of the study, and every student in this group had at least two Multistructural responses, it was however interesting to note that two of the Low group responses were assessed at the Relational level. So whilst there could be a relationship between the two areas it was also possible that the two skills were not inseparable, and that students who were themselves poor performers could hear clearly and express their thoughts coherently.

School Achievement

As the Pilot group contained 2 students who were outstanding school achievers and 2 who were outstandingly poor in all areas of school life (academic, social, sporting and musical), it was possible to isolate and compare the results of these two groups of students. The observed differences between the results of these two groups of students, as seen on Table 4.06, were striking, and seemed to give some support to the finding by Biggs and Collis (1982) that SOLO levels showed a close relationship with school achievement.

Table 4.06 Pilot Study Responses Grouped According to School Achievement

Groups	
Low	High
MUM	RRR
MUM	MRR

A preliminary investigation of the numbers of students in each grade who were outstanding achievers or failures in all areas of school life revealed that there were usually only 1 or 2 students in each category, and that only some of these students chose to study Music in Grades 9 and 10. So although it would have been possible to identify students for their overall school achievement in the Main Study, it was felt that the numbers would be too small to show any trends that could not be noted in the groups identified for other influencing factors. This characteristic was only noted here because this group of Grade 10s had contained these four exceptional students. The presence of these students also meant that this was not a typical Grade 10 music class, so that the results obtained from this group might differ from normal patterns, and this was another factor in the decision to conduct another study with a much larger sample of students.

Summary of Pilot Study Results

This examination of the results from the Pilot Study uncovered some interesting points, led to several interim conclusions, and indicated directions for further investigation. Although it seemed possible to analyse student responses using SOLO, as the Pilot Study group had been involved in intensive trials with listening materials throughout the year, it was not clear what level of responses could under more normal circumstances be expected from students in this age range.

From examination of the responses to the final test it seemed safe to conclude that students at this age and with this level of musical training were capable of producing responses from

Unistructural to Relational. No Prestructural responses were obtained, despite having been observed in responses by these students during the year. Although one response put forward some speculations, there was no evidence in these responses of original thoughts, indicating Extended Abstract tendencies, although these had been put forward by some of the students during the year. This may also have been due to shorter time being available for each piece in the test than that allowed during the Pilot Study, and to the pressure imposed by the end-of-year testing situation, where they were striving for maximum marks and deliberately avoiding any possible irrelevancies.

It was also uncertain whether response levels were constant, or whether they fluctuated according to the piece of music, the question asked or other unknown factors. The responses of four students had been rated at the same level for each piece, and the other nine students had equal levels for 2 pieces. It seemed likely therefore upon this occasion that the students were mostly functioning at the same SOLO level. The reason for the difference in the other level was not obvious. The piece by Varèse seemed to produce a wider spread of responses overall, and when examining individual response patterns, it could be seen that slightly more students (6) had a different response for this piece than the other pieces (4 and 1). Despite being the sole task which required a Relational level response, the Varèse stimulated more Unistructural level responses than the others, which might indicate that the students found this piece or question harder to handle. At this stage therefore it was not possible to determine whether or not the use of a different task, the selection of music containing disliked features, or some other unsuspected variable had been responsible for the slightly different pattern in responses to the second piece.

It could perhaps be asserted that the test was measuring Listening Ability as understood by the classroom music teacher. Some evidence for this assertion had been found in the comparison of SOLO assessments to the groupings of the students judged on their oral reponses to music in class (Listening Ability), but due to the small number of students and responses involved this was still a

supposition, and there was also some evidence of other influencing factors. There was also a possibility that the test could be measuring Listening Ability apart from other musical abilities as two students in the group identified as having the poorest Performance Ability skills were able to gain Relational levels, and 41% of the Average group also attained the Relational level, though this was the level of 89% of the High group's responses. It was therefore possible, though not common, for students to achieve at the highest level seen so far in this age group despite lack of performance skills.

It seemed that the patterns of achievement students had developed in school, held true for this test, with students who were outstanding in the academic, social and sporting aspects of school life achieving higher levels on the test than low school achievers. This might be an indication that the test was not one measuring listening ability alone, as in an audio test, but was measuring students' performance in a classroom, where all the normal school influences on performance such as writing ability and attitude would be present. The relationship between success at school and school-based activities such as writing, and motivation had been introduced and it was resolved to look at this in the Main Study.

Even though some questions seemed to have been resolved during the Pilot Study, the student sample had been very small and possibly not typical, so in order to be justified in drawing conclusions suitable for classroom implementation, it was considered necessary to carry out further investigations with a larger number of students. As a result of the Pilot Study it was decided to use the same test for the Main Study, but to make a few alterations to the test format, and to include Motivation towards Music lessons and Attitude towards the test in the investigation of influencing factors.

In the Pilot Study the test had been administered under test conditions and the responses had not shown as wide a spread of levels as when similar listening tests had been administered as part of a normal lesson. Accordingly changes to the test format were designed which, whilst retaining those features which ensured that each student's response was uninfluenced by other students, were intended to simulate a more normal classroom atmosphere.

Chapter Five

The Main Study

In the Main Study a total of 328 students from two Tasmanian High Schools were tested over a three year period. The schools were located in the same region of the state, were approximately of the same size, and had students from the same socio-economic background. The Music programmes were also similar with Concert Bands as the main performing groups. Both schools had two classroom music teachers and three itinerant instrumental teachers visiting for $2\frac{1}{2}$ days per week.

As the investigator conducted all the tests, availability of the investigator during the students' normal Music lesson time was the prime factor in the selection of students. Most of the students were in classes taught by the investigator but some other classes were made available through exchange of classes. The students tested were all in Grade 7, 8, 9 or 10, and were all taking Music as a subject. For the Grade 7 and 8 students it was a compulsory subject, but the Grade 9 and 10 students were in an elective course. The Grade 7 and 8 students were in homogeneous classes of 25-28 students and all came from the same school. The Grade 9 and 10 students came from both schools and had all chosen to study Music, so they could therefore be expected to have greater musical ability and interest in Music than the younger students. These elective Music classes were not homogeneous, and class sizes varied from 12-20 students.

Due to medical and counselling appointments 9 of the 328 students tested did not complete all three responses, and these response sets were discarded leaving a final total of 319 subjects. Two testing sessions were interrupted due to a fire evacuation and a power failure. Fortunately both interruptions occurred between pieces, and the tests were completed after intervals of 15 and 25 minutes. Responses from these test sessions were isolated and it was intended to discard them, but as, when these responses and response levels were examined they were found to be similar to those of parallel classes tested during the same weeks, these response sets were not discarded but included in the totals. The grades in which the tests were taken can be seen on Table 5.01.

Table 5.01 Students in the Main Study

	Grade in which Test was Taken				
Tests	Grade 7	Grade 8	Grade 9	Grade 10	Total
First Test	157	77	40	45	319
Second Test	0	48	15	27	90
Third Test	0	0	11	0	11
Totals	157	125	66	72	420

As testing was carried out over a two year period in the same school, it was possible to retest some students in consecutive years and gain some ideas of progress due to musical training or maturation. As classes were rearranged at the end of each year, retested classes did not always contain the same students and this accounts for the discrepancies between the numbers retested. As can be seen on Table 5.01, of the 319 subjects who completed tests, 229 students were tested only once, 90 were tested twice and of these 11 were tested for a third time. Of the students who took the test first in Grade 7, 48 were retested in Grade 8; 12 students were retested from Grade 8 to 9; 3 were tested twice in Grade 9; 14 from Grade 9 to 10; and 13 students were tested twice in Grade 10. The 11 students tested three times were in a Grade 9 class which took part in a short retest experiment, they had already been tested in Grade 8 and were tested twice in Grade 9 for the Retest Experiment. (This class consisted of 15 students, 3 students had not completed the test in Grade 8 and so have been listed as tested twice in Grade 9,

and one student, who was tested in Grade 8 but was absent for the Experiment's Pretest, is listed with the Grade 8 to 9 retest group. As the Retest Experiment was only concerned with changes from the Pretest to the Posttest it includes the responses of all 14 students who took the test twice in Grade 9.) Altogether 471 separate responses were collected from Grade 7s, 375 from Grade 8s, 198 from Grade 9s and 216 from Grade 10s, making a total of 420 completed response sets made up of 1260 separate responses.

Identification of Possible Influencing Factors

Students who were considered as being either outstandingly strong or weak on four factors which might have influenced their SOLO levels were identified. The four factors were : Listening Ability, Written Fluency, Motivation in Music Lessons and Music Performance Ability. So as to avoid any possible bias due to test familiarity, the data for these factors was all taken from the results of the students' first tests.

Listening Ability was defined as the ability to identify musical elements and talk about music heard in class. Students functioning at either extreme were identified by the Music subject teacher, who had observed their responses to class discussions about music listened to in class, and to questions put during composing and performing activities. As can be seen on Table 5.02 10 students were identified as having very poor skills in this area, and 16 as being outstandingly good. In the Low group, 3 came from Grade 7, 2 from Grade 8 and 4 from Grade 9. The High group included 11 Grade 7s, 3 Grade 8s and 2 Grade 10s.

Table 5.02 Students Identified for Listening Ability

	Listening Ability Groupings	
Grades	Low	High
Grade 7	3	11
Grade 8	3	3
Grade 9	4	0
Grade 10	0	2
Total	10	16

Written Fluency, the ability to express thoughts effectively on paper, is a skill ostensibly taught mostly in English lessons, so students' English teachers were approached to identify those whose skills in this area were outstandingly good or bad. As can be seen on Table 5.03, 23 students, who were mostly all having extra lessons on written expression with a special education teacher, were identified as the Low group. This group was made up of 14 Grade 7 students, 3 Grade 8s and 6 Grade 9s. The 19 students in the High group consisted of 10 Grade 7s, 3 Grade 8s, 1 Grade 9 and 5 Grade 10s.

Table 5.03 Students Identified for Written Fluency

	Written Fluency Groupings	
Grades	Low	High
Grade 7	14	10
Grade 8	3	3
Grade 9	6	1
Grade 10	0	5
Total	23	19

Motivation towards Music, defined as a combination of attitude towards the subject and willingness to complete tasks set, was assessed jointly by all the music teachers in each school. The Low group students were characterised by their failure to bring essential equipment (such as the music to be played) to classes, were slow to settle down, unwilling to cooperate with other students, and despite receiving extra help in lessons, still did not complete tasks set, and did not appear to be interested in doing so.

As Table 5.04 shows 17 Grade 7 students were in this group, 5 Grade 8s, 5 Grade 9s and 1 Grade 10 making a total of 28 students. The High group students had been identified because they were obviously keen, always worked hard in class, were eager to do extra work in Music, and were often loathe to leave and go to other subjects. Those that were technically able, also participated in extra musical activities outside their timetabled lessons. There were 16 Grade 7s, 5 Grade 8s, 4 Grade 9s and 6 Grade 10s, making a total of 31 in this group. Discussion of students in both of these groups with other teachers in the schools revealed that most of them showed similar behaviour patterns in other areas of school life. The Low group students were all problems in other areas as well, and most of those in Grade 9 and 10 were undertaking special Technical College Link programmes designed to provide a sense of purpose and some vocational skills for potential early school leavers. Although some of the High group were only enthusiastic and hardworking in Music, most of this group were successful in other subjects too.

Table 5.04 Students Identified for Motivation

	Motivation Groups	
Grades	Low	High
Grade 7	17	16
Grade 8	5	5
Grade 9	5	4
Grade 10	1	6
Totals	28	31

Music Performance Skill was defined as the ability to make music on an instrument. As all those students who showed talent were taught by the itinerant instrumental teachers, who were in a position to compare standards between schools, they identified the outstandingly good students in this area of the course. Any student who had great difficulty learning to play a simple tune accurately was placed in the Low group. There were 11 Grade 7s, 5 Grade 8s and 4 Grade 9s in the Low group of 20 students, and of the 24 students in the High group 11 came from Grade 7, 5 from Grade 8, 3 from Grade 9 and 5 from Grade 10. It had been thought initially that these groups might contain the same students as the Motivation groups, but

this was not the case. Many of the students observed as having a poor attitude towards Music lessons were better performers than some who were highly motivated and wanted to be involved with Music, but were not successful performers.

Table 5.05 Students Identified for Performance Ability

	Performance Ability Groupings	
Levels	Low	High
Grade 7	11	11
Grade 8	5	5
Grade 9	4	3
Grade 10	0	5
Total	20	24

Test Materials

Each student was provided with a separate A4 answer sheet for each piece of music, copies of which are in the Appendix. Name, age, grade, year and month were required to be written at the top, and as each piece was referred to by number, to remove any advantage given by recognition of the composer, period, form or programmatic title, there was also a box in which the student was asked to write the number of the piece to which they were listening. Lines were ruled halfway down the page to indicate the length of response expected, yet leave room for longer responses.

The answer sheets for pieces 1 and 3 were identical and had the listening task printed at the top :
Describe this piece of music as clearly as possible.

The second piece, Varèse, had a task which appeared to be different requiring the same information to be presented in a way that showed its relationship to the emotional mood created :
The composer of this music is trying to create a mood in your mind. In one sentence explain the scene or mood the music created for you, then ***explain why*** *the music creates this impression for you.*
As very few Pilot Study students had used the sentence indicating musical elements that could be included in a response, it was deleted from the answer paper. It was replaced by a pretest mini-lesson, with the aim of stimulating as many higher level responses as possible.

The music used was the same as that in the Pilot Study :
Elgar : Bavarian Dance No 3 Op. 27 "The Marksman" performed by the London Symphony Orchestra conducted by Lawrence Collingwood. The extract lasted 2 minutes, came from the second quarter of the dance and contained two contrasting sections with the first repeated in ABA form, and faded out with glockenspiel notes.
Varèse : Ecuatorial (1934) for bass, 4 trumpets, 4 trombones, piano, organ, 2 Ondes Martenot and 5 percussion performed by The Contemporary Chamber Ensemble directed by Arthur Weisberg. The extract used came from the opening and ended before the vocal line enters.
Telemann : Concerto in F for Recorder, Bassoon, strings and continuo performed by Michaela Petri, Klaus Thunemann with the Academy of St. Martin-in-the-Fields directed by Iona Brown. The extract is the first half of the second movement marked Vivace.

Test Venue

With one exception, which will be described later under the heading Moderation Experiment, all testing took place in the students' normal classroom during a timetabled lesson and was seen as being part of the normal half yearly assessment procedures. Each classroom was equipped with a good quality stereo cassette player, amplifier and large wall mounted speakers. Students were seated approximately one metre apart facing the speakers and main black-board.

Test Procedure

In order to make the test procedure as like a normal lesson as possible, to allay test anxiety and enable all students to produce responses closest to the optimal level in their Zone of Proximal Development (Vygotsky 1978), a mini-lesson on musical elements that could be used when talking about a piece of music was inserted into the test procedure before the pieces were played. This, revising the elements that make up a piece of music, ensured that all the students had the information needed to produce a Multistructural response, and as it was also stressed that all statements made should be proved by evidence taken directly from the piece of music heard and by reference to other pieces, it was also an attempt to promote the giving of Relational level responses. It also gave all students the opportunity to question and clarify basic concepts and, by ensuring that the less able readers and slow learner had heard the terms explained again close to the test, was also an attempt to compensate for the fact that this was basically a written test.

The mini-lesson was presented in two ways suitable to the level of their musical training, the Grade 7 and 8 students being encouraged to produce ideas in their own word from their previous experiences which were then expanded to form the complete list; whereas the older students were presented with the correct musical terms that they were learning and assisted to explain them.

Mini-Lesson for Grades 7 and 8

"Today you are going to hear and write about three pieces of music. Before we start let's go over the different aspects of the music which you could talk about in your answer..." A quick class brainstorm session followed with the students ideas clarified into the terms below and written on board. When list the was complete, with the teacher having added any missing aspects, a brief comprehension check of each term was carried out.

Grade 7 and 8 list :

Instruments	*Repeats of Tunes*	*Types of Tunes*
Contrasts	*Harmony/Chords*	*Speed*
Mood	*Volume*	*Rhythm and Accents*
Accompaniments		

Ways piece is Like other pieces of music you know

Ways piece is Different from other pieces of music that you know

The Grade 7 and 8 students were given more reminders and help than the older students in a deliberate attempt to redress the imbalance in amount of musical training between the two groups, and also to create a less threatening atmosphere.

Mini-lesson for Grades 9 and 10

The list below, which uses the musical terms the students had been learning during their course, was already written on the blackboard before they arrived. The list was already written down as it was felt in the amount of time available the terms produced from brainstorming would not be the ones appropriate for this level of tuition. The students were asked to contribute by explaining the meaning of each item on the list and giving examples from pieces previously used in class. Each item was explained by two different students and the teacher, then at the end of the list students were asked if they were now sure what each term meant, and any queries were answered by the teacher.

Grade 9 and 10 List :

Orchestration	*Timbre*	*Formal Structures*
Melodic Structures	*Genre/Style*	*Historical Perspective*
Texture	*Harmony and Tonality*	*Tempo*
Dynamics	*Rhythmic Impetus*	*Articulations*
Imagery		

Test Instructions

"First fill in the top of the sheet. This is the first piece so put 1 in the box. Put your pen down when you are have finished and when everyone is ready I'll tell you what to do next."

When all pens were down on desks and they were ready to listen :

"I am going to play you three different pieces of music and ask you to write about what you hear. You will hear each piece three times with gaps in between. You should have plenty of time to write but if you need extra time you can have it at the end after the third piece, or come back later. You can also hear the music again at the end if you need to listen to it again".

"Look at the question" .. investigator read it out .. *"Remember to prove whatever you say about the music with evidence from what you hear. When you have finished, put your pen down, and I'll know when to go onto the next piece".*

These instructions were repeated before each piece, with this addition after the question for the second piece :

"If the music doesn't create a scene of mood for you, don't worry, there aren't any marks for the scene or mood. All the marks are for explaining why the music creates this impresion, so if you don't get any impressions just write "no scene" on the top line, and go on and describe the music just as you did for the first piece."

In addition to the extra time given at the end of the test to check spellings and grammar, students were also told that they could have extra writing time and that they could also hear the pieces again if they wished. The intention behind the offering of extra time and hearings was to allay fears and provide a less stressful environment, not to actually provide more time or hearings. It was intended to restrict the extra time and hearings to those students who requested them, and then to discard those responses. This was not found to be necessary as no one asked to hear the music again before completing their responses, and only six students asked for extra time. These students each took between 2 and 4 minutes extra to complete their responses. Upon comparison of these responses with those from students who had completed their responses in the time allocated, it was seen that the responses were similar. It was

therefore considered that the extra time had not given these students any advantage, so their responses were retained rather than being discarded. It had also been expected that those requesting extra time would be the more able students, but with one exception, this was not the case. This exception was a capable academic student who did not write anything in the extra time, but used it to read through his responses carefully. The other five students used the time to complete their responses, which were all stories stimulated by the mood of the music, which were classified as Unistructural.

All students cooperated well with this procedure. Five students appeared to be ill at ease, so they were closely monitored during the test and given some positive reassurance in the form of smiles and brief comments such as *"That's right ... that looks good ...Well done ... Good. Keep going. You're doing well".* This seemed to encourage them to make more effort, and at the end they reported feeling happy with their responses. As, despite this extra encouragement, these responses were no better than others in the same classes, these responses were also not discarded as intended but were retained.

Student Feedback on the Test

At the end of each testing session an informal discussion was conducted during which students were asked about their feelings towards the three pieces. The piece most preferred and frequently requested to be played again in other lessons was the Telemann. Discussion was also heated over varying interpretations of the Varèse and these debates continued outside the classroom.

As the tests were conducted as part of the normal Music testing programme the responses were evaluated immediately and students provided with some feedback, though contrary to normal school practice the papers were not returned. Feedback, although based upon the SOLO levels, was designed to be positive, to increase the students' motivation towards any further tests, and it was given in

a form comparable to the normal school marking system. Two systems were used one using terms appropriate to Grades 7 and 8, and one suitable for Grades 9 and 10.

As it was intended to retest the students in Grade 7 and 8, it was felt important to develop a positive feeling towards the test and improve music listening skills, so positive symbols (which were already familiar from practical progress charts) rather than ratings were used for feedback with these grades. Alphabetical classlists with one symbol for each piece were posted on the wall together with the following explanation of each symbol (Figure 5.01). (The SOLO labels were not given to the students and are included here only to show the relationship between the feedback and the SOLO levels.)

Figure 5.01 Feedback Symbols and Explanations for Grades 7 and 8

*** *An excellent answer. You mentioned lots of aspects of the music and showed how they connected with each other. Well Done.*
(Relational)

* *A good answer. you mentioned many different things in the music.*
(Multistructural)

✓ *OK, but you could have mentioned more of the different things you heard in the music.*
(Unistructural)

? *You had some trouble, either you couldn't work out what to do, or you couldn't decide what it was you actually heard. Why not ask for help next time ?*
(Prestructural)

As most students received at least one *, the wall became a popular venue especially amongst low achievers who brought their friends from other classes to look at it.

A slightly different system of letter ratings was used for Grade 9 and 10 students (Figure 5.02), though once again the response sheets were not returned, and a ratings list was published

showing separate ratings for each piece. As the syllabus and teaching at this level were aimed towards developing the skill of explaining musical styles and idioms through the identification and connection of musical elements and the construction of the music, the ratings system was designed to reflect this.

Figure 5.02 Feedback Symbols and Explanations for Grades 9 and 10

A *An excellent answer. You mention many aspects of the music, explain your ideas clearly, and connect ideas and musical facts together efficiently.*

(Fully Relational)

B *You heard a lot of the ideas the composer put into this piece of music, and are starting to explain musical ideas using the evidence you have found in the music.*
Well done, you are becoming an intelligent listener.

(Multistructural + Relational ideas,
Multistructural Transitional response)

C *You have heard a lot of the ideas the composer put into this piece of music. Good.*

(Multistructural)

D *You only mentioned one aspect of the music. Didn't you hear the others ? or didn't you know how to write them down ?*
I think you may need some help.

(Unistructural)

? *You didn't answer the question.*
Did you understand what to listen for ?
You may need some help next time.

(Prestructural)

This system clearly indicates that a Unistructural response is unacceptable and that although a Multistructural response is acceptable, with a little more effort using the same material a higher award could be gained. The Multistructural response is seen as the basic standard from which to improve, and this seems to be reasonable as even the less able students had been observed to give responses at this level during the Pilot Study. Due to poor writing skills, responses from these students had been limited to short

phrases rather than complete sentences, but they heard more in the music than they wrote down, and could when motivated talk at length and write down extensive lists. Motivation to communicate ideas and poor communication skills had seemed to limit responses from these students, rather than the inability to hear, distinguish or recognise musical sounds.

Student reactions to these forms of feedback were favourable, and as other listening work was also assessed using this system, they were able to use the feedback to compare their own and other students' work. By using the wallchart explanation of the letters they were able to plan how to improve their work. [The TCE, which was introduced as testing ended, uses a similar letter rating system for assessment of its set criteria.]

Analysis of Responses

Responses were evaluated immediately after each test so that feedback could be given to the students. Some students responses were extensive and continued past the ruled lines and even onto the back of the page, but the majority were confined to the lines. Although some Relational responses were quite lengthy, quantity was not necessarily an indication of quality as the longest responses were those of the Unistructural storytellers. All responses were re-evaluated over a two week period at the end of the testing period to ensure that the same standards had been used throughout.

Although some Transitional responses were noticed, and those based in the Multistructural level but moving towards the Relational level were acknowledged in the feedback given to the Grade 9 and 10 students, they were not given a separate category. This was due to several factors. Within the Multistructural category, where there was a wide variety in the number of musical elements mentioned, a need was seen for several subcategories, for which reliable general identification criteria proved difficult to provide. Whilst some attempts at this were made and subclassifications were made of

responses from several of the earlier classes tested, as the information given by the subclassifications was more confusing and therefore less useful to students than that given by the SOLO categories, the value of this exercise was thought to be doubtful. It was also felt that as the purpose of the study was to establish whether the overall SOLO principles could be applied to Music Listening, it would be more useful in this first study to concentrate on the main SOLO categories.

A two part procedure was used for evaluation of responses. Firstly all musical elements mentioned correctly were recorded on the checklist developed in the Pilot Study (see Appendix F), then the response was allocated a SOLO level using the guidelines listed in Chapter 4 (Figure 4.02). A representative sample of the type of responses produced has been typed in italics below. All responses quoted include the original spelling, and as far as possible, the original layout. E indicates that the response was given to piece No 2, Elgar; V a response to piece No 2, Varèse; and T a response to piece No 3, Telemann. A fuller selection of responses given to each of the pieces of music is in Appendix J.

Prestructural Responses

Responses which did not attempt to answer the question were classified as Prestructural. As the question for pieces 1 and 3 was *"Describe the music as clearly as possible"* responses which did not mention the music at all were classified at this level. These responses all described scenes which had presumably been stimulated by the music, but as there was no indication of this they were classified as Prestructural rather than Unistructural.

"It is spring. My sister and I are in the garden as big as can be and I keep holdin on her then we see some squrel doing the same and the we see some birds doing the same then we all do it again." (T1, Grade 8)

The next response also fails to mention any aspect of the music, but the events within the story mirror the short episodes

within the music. This would be a satisfactory response to a creative writing task based on this music, but it is not adequate as a musical response.

"Two children - brother and sister run into a meadow to play. there is a hill the littel girl stops to pick a daisy and the little boy cartwheels down the hill, the girl runs behind.

At the bottom they play hide-and-seek in the long grass, then make daisy chains, the boy tires of this and throws grass at his sister she begins to chase her brother as they fall down laughing and play leap-frog until the little girl falls over and hurts her knee, and her brother comforts her and wipes away her shortlived tears they jump up to play TAG then race each other back up the hill to where their mother is waiting for them as they happily retreat indoors for lunch." (T2, Grade 9)

The second piece had a two part task where the second part was the one being evaluated. Responses that dealt only with the scene and did not "*explain why the music creates this impression*" were therefore not adequate answers to the question, and were classified as Prestructural. *"I had a picturer of a man running threw a street being followed by police or some monster and he ran into a church for cover.*

Or it could be an army man in the batel field walking alone and seeing dead boides lying allover the ground after a bomb hit, and he found his friend deas and turned him over on his back to look at his face." (V2, Grade 8)

Unistructural Responses

Responses which answered the question by using one aspect of the music were classified as Unistructural. Most of these responses mentioned speed, volume or the instruments which were playing.

"There is some one sneaking around where they are not supposed to be and nearly gets caught. After a while they decide to go somewhere else. The paino sounds like someones foot steps. the cornet sounds like another person looking for him the drums sound like a group of people on patroll." (V4, Grade 8 - instruments)

"This piece of music reminds me of two birds fighting over a

piece of bread both trying to grab it and munch into it. Then suddenly a huge crow comes and takes it from them both and the two birds fly away in disgust. Why I said two birds were fighting over it was because every little bit was repeated like a echo." (T3, Grade 10 - melody)

Multistructural Responses

73% of all responses were classified as Multistructural because they used several musical elements to answer the questions. There were differences in the quality of responses at this level, some being very simple whereas others showed a good grasp of musical elements, based their response logically around the structure of the piece and if they had not resorted to using lists would have been classified as Relational.

This response does not show as much care as some of the Unistructural responses, but does show awareness of more aspects of the music and so was classified as Multistructural;

"piece of music of mood. Volume soft to loud. A bit of speed. Getting faster." (E3 Grade 8)

There were also more musically descriptive responses that were presented with more care, such as this one.

"very quick going up and down the scales very fast with the viloin but not so fast with the tuba always changeing very exciting but it makes you feel relaxed more instruments come in along the way it repeats parts over and over again it is very jerky" (T4, Grade 7 - melody, instrumentation, mood, form, rhythm)

The most common type of Multistructural responses was longer and more detailed like this one :

"It is harmony of bass instruments and woodwind. The dynamics change constantly, there is a big build up and then it goes softer. It is rather fast, has a jumpy melody, quite fast. There are lots of different instruments playing at the same time. Starts off at a reasonable pace gets faster and louder. It's like both Brass and Woodwind are competing against each other. Some parts it is tongued

others it is slurred." (E7, Grade 9 -instrumentation, volume, melody, speed, articulation)

There were also some competent Multistructural responses that were mostly logically based around the formal structure of the pieces and seemed to be almost ready to make the transition to the Relational level. They still seemed to be making lists, so they were classified as Multistructural :

"This piece of music is VERY fast, it's lively, and it jumps about alot, the violins have the tune mostly, but after a while a horn or what sounds like a horn but I'm not sure if it is a horn, takes the tunes and going on a bit, and after that a recorder, or what sounds like one, has the tune, it has a harmony line, and different instruments play that, so it varies it tunes and harmony lines. There is also repeats of the beginning, and of different parts in this particular piece of music. I think the composer must have had a hard time writing this music onto paper it must've taken hours too do. It has lots of trills in it, not exactly hundreds but quite a lot. Overall, this piece was pretty good to listen too." (T7, Grade 7)

Relational Responses

Although all Relational responses created logical arguments by basing their description on one Group Two element such as form or style and using other elements as illustrations or evidence for their thought, these responses also varied in their quality. Each of the responses at this level was based on one aspect of the music, usually the formal structure of the piece, and the difference in quality lay in the way other elements were integrated.

"The beginning of the piece is hesitant and detached. It forms an image in my mind of a criminal lurking around a deserted alleyway, constantly watching out for the police.

The detached notes at the beginning and then the long high pitched notes indicate the criminal tramping around the alleyway, but still being observant, looking out for the police.

The sudden crash of the bass drums and then the increase in speed and volume indicate the criminal has been spotted, and he

begins, desperately, trying to run along the cold deserted alleyways to keep in front of the police. As the piece reaches a descrescendo at the end, the effect of the gradual softness indicates the criminal has managed to disappear into the distance, never to be seen again..... " (V10, Grade 10)

Relational responses, which integrated many musical elements effectively to form a coherent argument, were usually found in Grade 10 responses, however they did also occur in Grade 7 responses.

"The music creates a dark, morbid atmosphere which creates a feeling of fear and suspense; perhaps ultimately death. The two main contributing factors to this atmosphere are probably the selection of instruments and the abstract, unpredictable tune which doesn't give the slightest clue to what's going to happen next. The use of the piano and trumpet at the start with the detached tune and minor feel set the atmosphere which is added to by the occasional, sometimes unexpected percussion. The section where the trumpets crescendo and stop suddenly gives the impression of a climax to some event. The music also tends to create a graveyard type, supernatural effect which is added to by the use of a whistle or a flute to create a owl's hoot. The organ also seems to support this idea as it is very like an organ church which is often used for this very effect." (V11, Grade 10)

Although for the purposes of this initial study it has been resolved to classify responses into the broad SOLO categories, during the assessment process transitional responses were noted, and it has been noted that a finer classification of responses could be possible.

Independent Judges

As in the Pilot Study, the reliability of the classification of responses was checked by the reassessment of a representative sample. Six experienced secondary music teachers (including the two who had assessed the Pilot Study sample) were asked to assess a sample of 12 responses. Each response came from a different student, there were 3 responses from each grade, 4 from each piece and as far as possible they came from each SOLO level, so within the limitations of the number which the judges felt they could handle in the limited time available, the sample was representative of the variety of responses which had been collected. To avoid any possible prejudices on the part of the judges and ensure that judgements were made on purely musical grounds each response was typed with spelling and grammar corrected, and was presented anonymously. Responses were first assessed according to the teacher's normal method and then ranked within each piece, after which the SOLO principles were explained and the guidelines were produced. Each response was then classified into the appropriate SOLO level and in every case the SOLO levels mirrored the rankings made based on their normal assessment practices. Every response was given the same SOLO level by each judge. Samples from 18 different students were also assessed in a similar way with the same result by three other music teachers in the Moderation Experiment.

Moderation Experiment

In November 1988 an attempt was made to use the test as a moderating instrument between the 28 students in two parallel Grade 10 Music classes in the same school. One of the classes had already taken the test six months earlier. Both classes moved into the larger classroom and took the test together (this was the only class not tested in their normal surroundings) and all 84 responses were evaluated using the SOLO assessment procedure.

A sample of 18 responses from different students, 6 selected from each piece, were selected to represent all the SOLO levels present. The responses were then corrected (spelling and grammar

only), typed and numbered to avoid any possible bias due to handwriting, presentation or expectations caused by the student's name. The typed responses were given to 4 experienced teachers who, apart from the investigator, had not previously been involved with SOLO. As in the previous judging with 6 other teachers, responses were typed and numbered rather than named, and the teachers were asked to assess and rank responses using both their normal marking scheme and SOLO levels.

Although the four judges classified responses into the same SOLO levels, there was some disagreement about the ranking of the middle range students using their normal marking and ranking schemes. This seemed to be because some teachers used presentation and quantity of ideas put forward as the basis for assessments within the Multistructural range and did not give much credit for depth of thought unless it was obvious throughout the entire response. Some of the responses for which this disagreement in the rankings based on the normal marking schemes occurred were Multistructural ones which showed some signs of being in a state of transition towards the Relational level, and it seemed that these teachers recognised the superior quality of the Relational responses and took account of this in their normal marking schemes, but did not seem to recognise that other responses might contain the same quality of ideas.

The Tasmanian Schools Board courses at this time were arranged in three ability levels. Level 1 courses were designed for the lower ability students, Level 2 for average students and Level 3 for the higher ability students. (Level 3 courses in Grades 11 and 12 were tertiary entrance requirements.) Despite the difference in standards between the levels, and some minor differences in course content, Music students were commonly taught in one composite class containing students working at all levels. This gave students the opportunity to work beside students achieving at higher levels, and sometimes stimulated them to produce a high standard of work in some areas of the course. As the Level 3 syllabus stressed the skill of linking ideas and concepts to each other, which could be held to be the basic principle behind the SOLO Relational classification, there

was an expectation that students working on the Level 3 syllabus would produce responses at the Relational level

Although the assessment and rankings of the middle range students which had been made using the judges own assessment methods did not always agree, SOLO levels were consistent between teachers. It was also noted that despite the fact that the School Board Levels given to the students during the course were decided on the basis of work in all three areas of the Music course : Performance, Composition and Listening, the patterns of the response sets mirrored the Schools Board levels almost exactly, with the majority of Relational responses being produced by the Level Three students. As the Level Three syllabus stressed skills and concepts which involved linking ideas, which could be held to involve thinking at a Relational level, this trend was the one which might have been anticipated. As this Moderation Experiment seemed to have indicated that there might be a link between the SOLO levels and course assessments, the Schools Board levels awarded to each of the Grade 9 and 10 students in the study at the end of their Grade 10 year were collected for comparison with their SOLO levels.

Retest Experiment

If SOLO was to be really useful in the classroom as a formative tool, it should have a role in the teaching process as well as being an evaluation tool. In November 1988 a short experiment was conducted with a group of Grade 9 students in an attempt to determine whether it was possible to teach students to improve their SOLO levels. 11 of these students had completed the test in Grade 8, and the other 3 were new enrolments into the school.

All 14 students were tested at the beginning of November and were told that they would be tested again with the same test during their end-of-year examination period at the end of the month. They were motivated to work hard by being told that the improvement in scores between the two tests would be used in their end of year assessment. In the three week interim, once a week they were given a

20 minute listening lesson during which the principles behind the SOLO assessment were explained, and then used in a discussion of the relationships between elements in a piece of music. A short written task was also set in each lesson. Appendix L contains further details of these lessons. The written task set was marked immediately using the 9/10 feedback letter rating shown in Figure 5.02, and the application of the marking system to each piece of written work was individually explained when requested. They were not however given any information or feedback from the first tests.

The class were also at this time working on individual compositions and were preparing individual instrumental performances, so they were therefore involved intensively in work-in-progress listening and criticism. This work was also discussed during this time in terms of the way the manipulation of musical elements could create changes in mood and character of the piece.

As will be seen in Chapter 6 the only students whose SOLO levels did not rise on the posttest were those whose original responses had all been at the Relational level. The biggest improvements came from four students with pretest responses at the Prestructural and Unistructural levels whose posttest levels were Multistructural on all three pieces. It seemed to be possible either to teach students how to respond at a Multistructural or a Relational level, or to motivate them to achieve their optimum level. This was a very short experiment and it was unfortunately not possible to retest these students the next year to possibly ascertain whether the improvement had been due to short term motivation, test familiarity, the tuition or some other unconsidered variable.

Summary

In this part of the study 328 students were tested, though due to incomplete response sets for various reasons, the responses from only 319 students were retained for analysis purposes. Retesting was carried out on 90 students after intervals of a year, six months and

for 14 Grade 9 students after three weeks as Retest Experiment. The 11 students who were tested three times came from this Retest group.

Responses were assessed by the investigator, in class groups immediately after each test, and later when all testing had finished every response was reassessed. Responses were also assessed by 9 experienced secondary music teachers in an evaluation of a representative sample of responses and as part of a Moderation Experiment between two Grade 10 classes in the same school. These teachers used their normal assessment procedures to rank responses, and then reassessed responses using the SOLO guidelines (Figure 4.02), although ratings using their own assessment methods varied in some cases all judges gave the same SOLO levels.

In an attempt to assess the influence of musical and non musical factors upon the test levels, students who were outstandingly strong or weak for four factors, Listening Ability, Written Fluency, Motivation towards Music lessons and Performance Ability were identified. The levels of responses from these students were later compared to those of students not so identified.

The value of the SOLO test as an evaluation procedure for the elective Music classes in Grades 9 and 10 was examined through a brief Moderation Experiment, and an examination of the Schools Board Levels awarded at the end of the two year Schools Board Music course. A brief Retest Experiment was also carried out with a Grade 9 class, in an attempt to determine whether specific training could improve performance on the test.

Chapter Six

Analysis of the Data Collected from the Main Study

In this chapter data from the Main Study will be grouped in six sections based around the major hypotheses outlined at the end of Chapter Four, the Pilot Study. These six sections are set out below.

1) Firstly the overall pattern of responses from all students tested will be presented to show that responses to different pieces of music from students in Grades 7 to 10 can be evaluated using the same SOLO technique.

Hypotheses :

H_o 1.1 The SOLO Taxonomy cannot be used to evaluate Music Listening responses.

H_o 1.2 Open ended questions will not stimulate responses at all SOLO levels.

2) An attempt to ascertain whether the three tasks set were equivalent will be made through the comparison and analysis of responses to each separate piece. Responses from each grade and individual student response patterns will also be taken into consideration here. Also if, as observed in the Pilot Study, students were reluctant to make an effort for a piece of music which they disliked, then responses to the second piece, which contained unpopular musical ideas would have lower SOLO levels than responses to the other pieces.

Hypotheses :

H_o 2.1 Different questions will have no effect upon SOLO levels.

H_o 2.2 Different styles of music will have no effect upon SOLO levels.

H_o 2.3 The selection of pieces containing disliked musical ideas will have no effect upon SOLO levels.

3) The question of the possible effect of test familiarity, maturation and musical training upon SOLO levels will be investigated through examination of responses from those students who were retested. This will include the responses from first, second and third tests; all first tests separated into grades; the first and second responses of students retested separated into grades; and the Grade 9 Retest Experiment.

Hypotheses :

H_o 3.1 Test familiarity will have no effect upon SOLO levels.

H_o 3.2 Increased age will have no effect upon SOLO levels.

H_o 3.3 Length of school musical training will have no effect upon SOLO levels.

4) Some other possible influences on student SOLO levels both musical and nonmusical will also be considered, through the comparison of responses given by those students identified as outstanding for Music Listening Ability, Written Fluency, Motivation in Music Lessons and Musical Performance Ability. Responses of these students will be compared to each other and to all other responses. In addition an examination of these students' individual response set patterns and those of students with at least one Prestructural or Unistructural or Relational response will be carried out to discover any similarities or differences not revealed in the general data.

Hypotheses :

Ho 4.1 Music Listening Ability will have no effect upon SOLO levels.

Ho 4.2 Written Fluency will have no effect upon SOLO levels.

Ho 4.3 Motivation in Music Lessons will have no effect upon SOLO levels.

Ho 4.4 Music Performance Ability will have no effect upon SOLO levels.

5) As the initial reason for the study was to devise an effective assessment technique for Grade 9 and 10 Schools Board courses a comparison was made between SOLO levels on this test and the three levels awarded by the Schools Board of Tasmania for the Music School Certificate course. These awards made at the end of the two year course included assessments in all three areas of Music : Composition, Performance and Listening. Data from the use of this test as part of a moderating instrument to identify level 3 students in two parallel Grade 10 classes will also be presented.

Hypotheses :

Ho 5.1 The Schools Board Music Course Levels will show no relationship with SOLO levels.

Ho 5.2 SOLO levels cannot be used as an indicator of Level Three standard in the Listening area of the Schools Board Music Courses.

6) The final section deals with an analysis of the musical elements used in responses at all levels, in order to determine whether the observation made in the Pilot Study that Relational responses seemed to make greater use of Group Two concepts was justified. The Groups of musical elements referred to in these hypotheses are those which were developed with the assistance of practising music teachers in the Pilot Study, and were listed on Table 4.04.

Hypotheses :

Ho 6.1 Musical elements cannot be classified into groups based on SOLO levels.

Ho 6.2 SOLO levels will show no relationship with the groups from which the elements within it were taken.

As SOLO classifications are not an interval ranking but fall into a nominal, or imprecise ordinal scale, nonparametric statistical testing procedures were used throughout in the analysis of data. For clarity, percentages rather than actual data will be used whenever uneven groups are to be compared, unless the group is too small to make this a legitimate procedure.

Total Data from the Main Study

The data examined in this chapter came from the Main Study in which 328 students were tested. Of these, 9 students were unable to complete all three questions, and responses from these students were therefore discarded. This left a total of 319 students in the sample.

As was seen on Table 5.01 in the previous chapter, at the time of their first test 157 students were in Grade 7, 77 in Grade 8, 40 in Grade 9 and 45 in Grade 10. Of these, 229 students were tested only once, 90 were tested twice and 11 of these students were tested again for a third time. Of the 90 students retested, 48 students first tested in Grade 7 were retested in Grade 8; 12 students were retested from Grade 8 to 9 and as 3 students had both their first and second tests in Grade 9 (in the Retest Experiment), there was a total of 15 second tests in Grade 9. There were 14 students retested from Grade 9 to Grade 10, and 13 students took the test twice in Grade 10. Of the 14 students who were retested in Grade 9 for the Retest Experiment, 11 had already been tested previously in Grade 8, and this accounts for the 11 students who were tested three times. Altogether 157 response sets were collected from Grade 7s, 125 from Grade 8s, 66 from Grade 9s and 72 from Grade 10s, giving a total of 420 completed tests. As each test contained 3 separate responses, this made a total of 1260 individual responses.

Altogether 157 response sets were collected from Grade 7s, 125 from Grade 8s, 66 from Grade 9s and 72 from Grade 10s, giving a total of 420 completed tests. As each test contained 3 separate

responses, this made a total of 1260 individual responses. All these responses were classified into SOLO levels twice by the investigator, and samples were classified by 9 experienced music teachers on two separate occasions, which would seem to refute H_o 1.1, that the SOLO Taxonomy cannot be used to evaluate Music Listening responses.

Students in this study tended to produce mostly Multistructural responses and 72.62% of all responses were classified at the Multistructural level. As can be seen in the totals on Table 6.01, of the remaining responses, 15.64% were at the Relational level, 9045% at the Unistructural level and 2.31% at the Prestructural level. Differences significant beyond the 0.001 level between overall response levels for each grade were found.

Table 6.01 Classification of All Responses as a Percentage of the Total for each Grade

	Grades of Students in Study				
Levels	Grade 7	Grade 8	Grade 9	Grade 10	Total
Prestructural	1.07	4.54	3.45	0	2.31
Unistructural	9.35	13.07	9.60	3.24	9.45
Multistructural	86.42	79.74	49.50	51.39	72.62
Relational	3.19	2.67	37.38	45.37	15.64

With Prestructural and Unistructural categories combined to remove the empty cell, $X^2=336.24$ at df=6 significant at 0.001 level.

There were minor differences between responses in Grades 7 and 8, but the major feature of the differences between grades was the marked increase in the percentage of Relational responses in Grades 9 and 10, accompanied by corresponding decreases in Prestructural and Unistructural responses in these grades. A X^2 re-examination of the data for the U,M and R responses only, clearly showed the main source of the significant differences to be the Relational level, as the X^2 calculated for each grade at this level exceeded 39 and $X^2=22.6$ demonstrated significance at the 0.001 level with df=6. These differences can be more clearly seen in Figure 6.01, which shows the response levels from each grade in diagrammatic form. (On Figure 6.01 and all other figures in this chapter the Prestructural

responses have been omitted, because the numbers for each grade were not only very small but also very similar.)

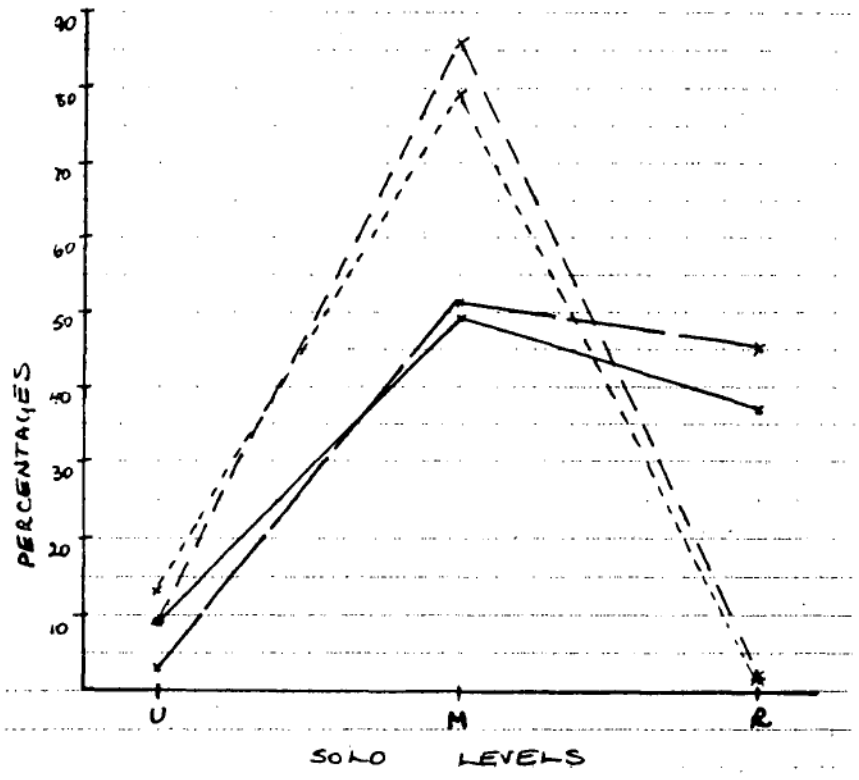
The diagram on Figure 6.01 shows quite clearly that the Multistructural and Relational responses from the Grade 7 and 8 students seemed to follow a different trend to that shown by the Grade 9 and 10 students. As all the Grade 9 and 10 students tested were in elective Music classes it was not certain whether this trend towards an increase in the proportion of Relational responses at this level was age related and due to greater experience; was an effect of the more intensive listening undertaken as part of their more specialised Music courses; was perhaps due to the greater interest in Music shown by their decision to select it as an elective subject; or was due to some other variable. As the data presented later from the Retest Experiment will show the strong effect upon SOLO levels of specific tuition, it will be seen that these sets of data could be used as initial evidence to support the hypothesis that tuition can affect SOLO levels on this test.

Response Patterns between Individual Pieces

Response patterns between individual pieces were examined closely to determine whether reactions to the different styles of music or the two types of question had assisted or prevented students from attaining their optimal level of response, and also to provide evidence upon which to judge H_o 2.1, 2.2 and 2.3, which are all concerned with the possible effects of different questions and pieces of music.

Responses for each piece varied both overall and within each Grade group, with the overall pattern replicating that found in the Pilot Study. Table 6.02 shows that whereas the majority of responses for each piece were at the Multistructural level, the majority of Prestructural and Unistructural responses were for the second piece, Varèse, which also stimulated a slightly higher proportion of the

Figure 6.01 Comparison of U, M and R Responses from each Grade



Relational responses.

Table 6.02 Classification of All Responses by Pieces
(presented as a percentage of all responses to each piece)

	Pieces of Music		
Levels	Elgar	Varèse	Telemann
Prestructural	0.005	5.96	0.72
Unistructural	2.15	20.15	5.96
Multistructural	82.15	57.38	78.34
Relational	15.48	16.43	15

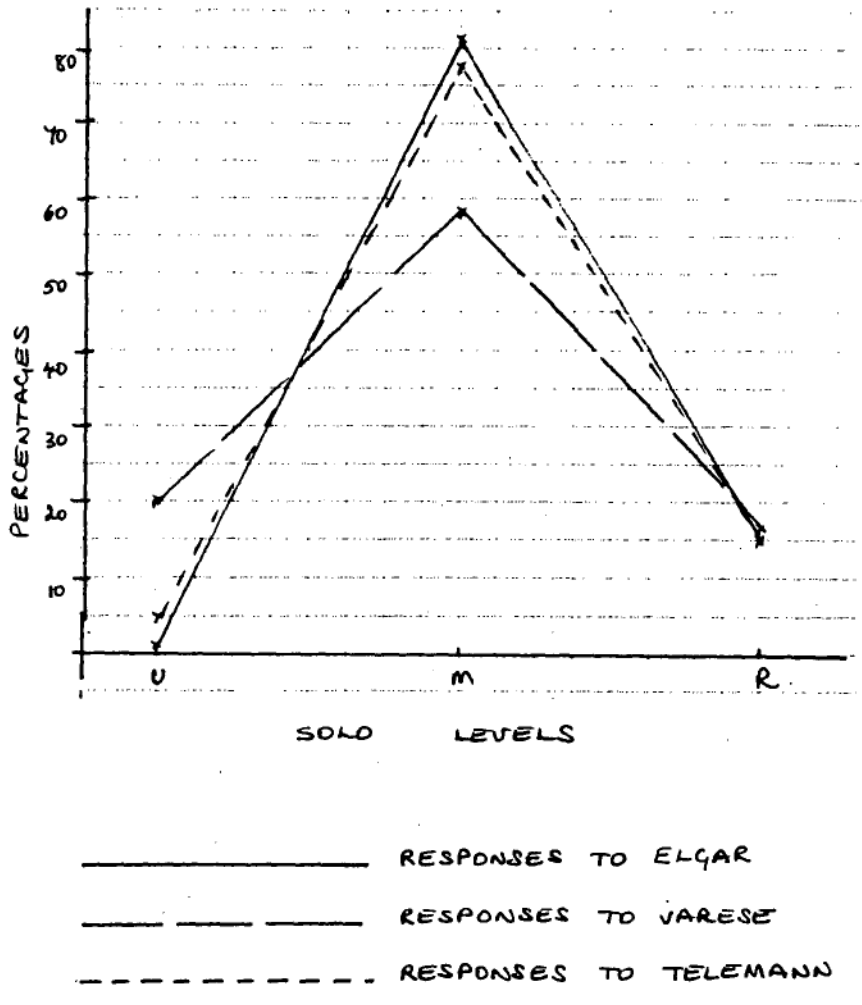
X^2=129.38 at df=6 significant at 0.001 level

The significance of the observed differences between the pieces was confirmed by use of the X^2 Test. The calculated value of X^2 for this data was calculated as being 129.38, which is far larger than the value of 22.46 for df=6 given for the 0.001 probability level. The null hypothesis that there was no difference between the results on the three pieces of music was therefore rejected, and the alternative hypothesis that there was a difference between responses on the three pieces was accepted. The next step seemed to be to identify the source of this difference.

Observations of the data on Table 6.02 had seemed to indicate that responses to the Varèse had been the source of the differences, and this was confimed visually by a diagrammatic comparison of responses from each of the three pieces on Figure 6.02. It can be seen on this figure that, although there were almost identical numbers of Relational level responses, there were marked differences at the other levels. The pattern of responses for the Elgar and Telemann pieces were very similar, but the Varèse stimulated more Unisructural responses and had a correspondingly lower number of Multistructural responses.

An analysis of the response levels for each grade was carried out to ascertain whether this pattern of responses was common to all grades. Details of this analysis are in Appendix K, and are summarised here.

Figure 6.02 Comparison of U,M and R Responses for Each Piece



In Grade 7, 8 and 9 the differences between the pieces were significant and the source of the differences was the Varèse. This had been expected as this piece had the only question demanding a Relational level response. However, as the greatest differences were not between the numbers of Relational responses for each piece as would be expected, but between the Prestructural and Unistructural responses, the possibility exists that these differences were not caused by the level of response demanded by the question, but may have been caused by students having difficulty understanding the requirements of the question.

There were also differences between the proportion of responses at each level within grades. There were for instance, no Prestructural responses from the Grade 10 students, which might indicate an increase in listening skills of these students due to the greater length of their musical training; yet the highest proportion of this level of response occurred in Grade 8 rather than Grade 7, which was the grade with the shortest amount of musical training in this area. Unistructural responses were also scarce in Grade 10, but Multistructural responses formed the majority of responses for all pieces in all grades, perhaps indicating that this is the level to be expected of the majority of students of these ages. Perhaps the difference in the proportion of Relational responses between the grades is the most interesting feature of this data. It seems to be the main difference between responses in the two lower and two higher grades, accounting for 3% of the total in Grades 7 and 8, and 41% of responses in Grades 9 and 10. Perhaps a closer examination of response patterns through individual responses might reveal more.

Individual Response Patterns

Although Biggs and Collis (1982) have stated that students may perform at different levels, 248 of the 420 sets of responses, comprising 59% of all response sets, were classified as having the same SOLO level for all three pieces. 1 student had all three responses as Unistructural (UUU), 208 had 3 Multistructural responses (MMM), and 39 had all responses at the Relational level (RRR). Table 6.03 also shows that 38.34% of response sets had two levels the same and only 2.62% had three different levels. When the responses to the Elgar and Telemann pieces, which had exactly the same question, are isolated, the number of response sets with 2 levels the same rises to 362 or 86.19% of all response sets, and 58 response sets, or 13.81% of the total, had 2 different levels, which confirms the pattern noted on Figure 6.02.

Table 6.03 Individual Response Patterns

	Similarities in Response Levels			
Main Levels	3 Same	2 Same	All Different	
Prestructural	0	2	-	
Unistructural	1	13	-	
Multistructural	208	126	-	
Relational	39	20	-	
No main level	-	-	11	
Total Responses	248	161	11	420
Totals as Percentages	59%	38.34	2.62	

On Table 6.03 the 2.62% of response sets having three different levels consisted of 11 response sets which fell into two groups. One group consisted of 8 response sets having Prestructural, Unistructural and Multistructural responses (1 MUP, 7 MPU) and the other having three response sets with Unistructural, Multistructural and Relational level responses (1 MUR, 2 RUM). In all but one of these response sets (MUP), the lowest level was for the second piece, Varèse. If these 11 three-level response sets are omitted, the remaining 409 response sets can be sorted into four groups according to the majority of their responses as has been done on

Table 6.04.

Table 6.04 Response Pattern Distribution of the Responses with 2 or More Responses at the Same Level (97.4% of total responses)

	Individual Response Patterns					
Main Level					Total	Total %
P		PPU 1	MPP 1		2	0.48
U	UUU 1	UPU 2	UUP 1			
		UMU 3	MUU 6	UUM 1	14	3.34
M	MMM 208	MPM 14				
		UMM 1	MUM 72	MMU 6		
		MRM 17	MMR 4	RMM 12	334	79.53
R	RRR 39	RMR 6	RRM 3	MRR 11	59	14.05

It can also be noted from Table 6.04 that MMM, all responses at the Multistructural level, was the most common pattern, obtained by 208 students. MUM, Multistructural with a Unistructural response for the second piece was the next most popular, achieved by 72 students, and RRR, all responses at the Relational level was attained by 39 students. The distribution of the Relational response sets as seen on Table 6.05 was also interesting, with all 39 RRR sets being found in Grades 9 and 10.

Table 6.05 Distribution of Response Sets containing R Responses

	Grades of Students				
Levels	Grade 7	Grade 8	Grade 9	Grade 10	Total
R	11	9	35	40	95
RR	4	0	23	32	59
RRR	0	0	17	22	39
R+% of Grade	1.67	1.72	12.64	13.24	

[R means 1 or more Relational response.

RR indicates 2 or more Relational responses.

RRR shows that all three responses were at the Relational level.

RR includes all RRR sets, R+ included both RRR and RR sets.

R+% of Grade gives the percentage of response sets in that Grade which included at least one Relational response, this is not the total percentage of individual Relational responses in the grade but the number of response sets that included responses classified as Relational.]

As can be seen on Table 6.05 Relational level responses were found in each Grade group, but were more concentrated in the higher Grades. In Grades 7 and 8 no response sets had all three responses at the Relational level, only 4 students had more than one Relational response and less than 2% of the response sets in Grade 7 Grade 8 included even 1 Relational level response. This contrasts with the Grade 9 and 10 pattern, where 13% of Grade 9 and Grade 10 response sets included at least one Relational level, and response sets with three Relational responses were also found, accounting for 9% of the total response sets. It may be that an examination of the responses of the retested students may throw more light on these patterns.

Retest Data

The data obtained from students who were retested is shown on Table 6.06. Although the number of students retested is not large, 90 students tested twice and 11 students tested three times, interesting trends do seem to emerge from the data.

Table 6.06

Classifications of Responses on All Tests as Percentages

	Tests		
Levels	First	Second	Third
Prestructural	1.78	4.45	0
Unistructural	10.24	7.78	0
Multistructural	77.43	59.63	42.42
Relational	10.56	28.15	57.57
[No of students	319	90	11]

Between First and Second Tests $X^2=35.59$ at df=3 significant at 0.001 level.

For first and second tests most responses were at the Multistructural level, with the Relational level having the next highest percentage of responses, followed by Unistructural, with Prestructural having the fewest responses. Although this order was

followed in both tests, in the first tests there was only 0.34% difference between the Relational and Unistructural levels, though the gap was much wider, 20%, in the retests. There was also a large increase in the percentage of Relational level responses from 11% on the first tests to 28% on the retests, a small increase in responses at the lowest level Prestructural responses and corresponding decreases in Unistructural and Multistructural responses. The differences on Table 6.06 between the first and second tests were found to be significant beyond the 0.001 level using the X^2 Test, and can be seen quite clearly on Figure 6.03.

Figure 6.03 clearly indicates the trend observed on the second tests towards an increase in the number of Relational responses. The patterns seen on this figure for First Tests and all the tests are very similar, whereas the Second Test pattern is noticeably different at the Multistructural and Relational levels. The percentage of Relational responses obtained on the second tests, 28.15%, was higher than that obtained from all tests, 15.64%, and much higher than the percentage obtained from Grade 7 and 8 responses, 3.19 and 2.67%. It was however not as high as those obtained from Grade 9 and 10 students, 37.38 and 45.37% (Table 6.01 and Figure 6.01).

These trends seem to provide evidence to refute the third group of hypotheses:

3.1 Test familiarity will have no effect upon SOLO levels.

3.2 Increased age will have no effect upon SOLO levels.

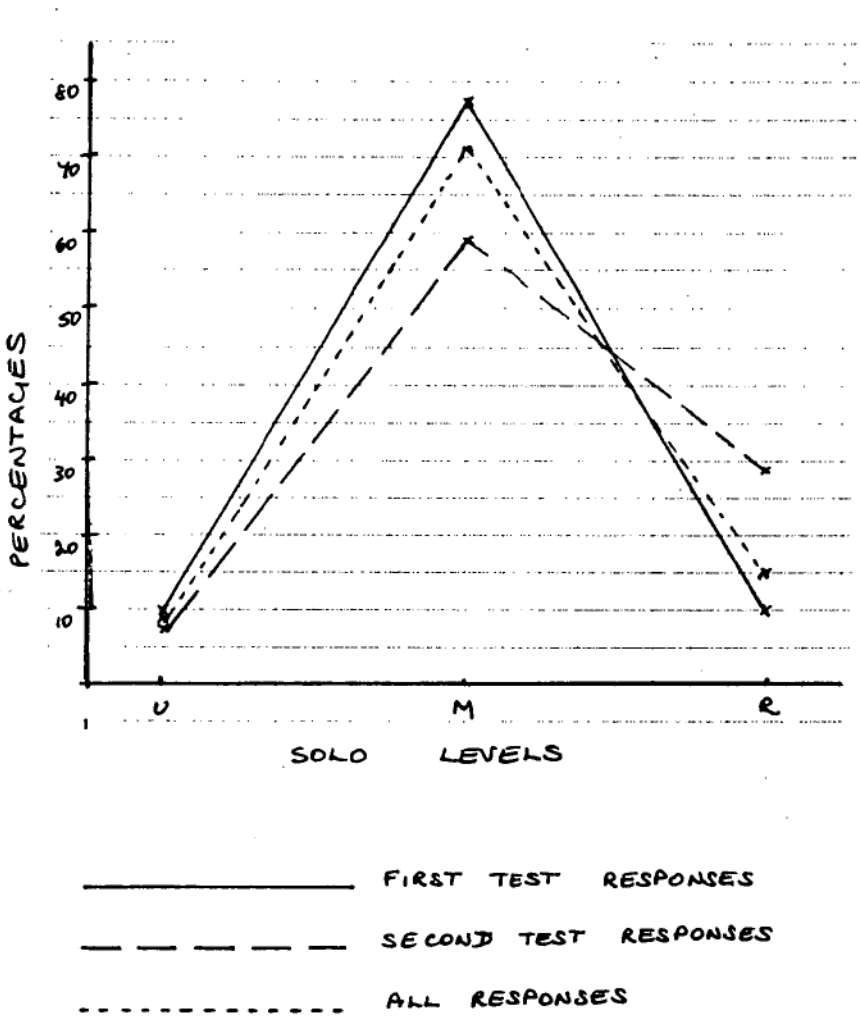
3.3 Length of school musical training will have no effect upon SOLO levels.

However these differences could be due to several factors including each of those above, and combinations of these and other factors. The higher percentage of Relational level responses on the second tests for instance, could be attributed to test familiarity and a change in the proportion of students from each grade in the second tests.

Of the 90 students who took the test for the second time, none came from Grade 7, 48 came from Grade 8, 15 from Grade 9 and 27 came

Figure 6.03

Comparison of U,M and R Responses on First and Second Tests



from Grade 10. As can be seen from Table 6.07, the relative proportions of students in each grade is completely different from the first tests, where the largest percentage of responses came from Grade 7 students. It could therefore also be argued that the increased age and therefore increased musical training of the second test students was an important factor in the production of a greater number of Relational level responses.

Table 6.07 First and Second Test Students as Percentages of Total Tests in each Test

	Grades			
Tests	Grade 7	Grade 8	Grade 9	Grade 10
First Test	49.22	24.14	12.54	14.11
Second Test	0	53.34	16.67	30

In addition, as the greatest proportion of Relational responses came from the Grade 9 and 10 students in the elective Music classes, (shown on Figure 6.01) it could be argued that the increased amount of musical training experienced by these students was a major factor, though for these students increased age was also a factor. The students who took the test three times were perhaps the only group for whom musical training could possibly be isolated as a major factor, and on Table 6.10 it can be seen that these students produced no Prestructural or Unistructural responses, and 58% of these responses were at the Relational level, a definite change in pattern from the other tests. It must be remembered that only 11 students were tested three times, and that not only were they the only students in the study who understood the SOLO principles upon which their responses were to be assessed, but they were also given strong motivation to produce high level responses. It is therefore inadvisable to credit these third test results solely to motivation, knowledge of the SOLO principles, or to the tuition given as the treatment. The sample of students in this group is also too small, and could be unrepresentative of the general population in Grade 9 Music classes, for this data to be used as anything other than an indication of possible trends. The differences between these tests and the third tests were not examined statistically as they are obvious, and as the third tests were not obtained under the same

conditions as the other tests.

As the possibility exists that the differences observed on the repeated tests could be due to test familiarity rather than to musical experience or training, an examination of the first test response levels separated into grades was made and is shown on Table 6.08. The X^2 test showed that significant differences existed between all grades. Differences significant at the 0.01 level were found between first tests in Grades 7 and 8, and Grades 9 and 10, and as the music courses undertaken by each of these pairs was similar, this difference could possibly be attributed to the effect of the one year difference in the amount of musical training and musical experience between the grades.

Table 6.08 First Test Results by Grades as Percentages of Responses from each Grade

	Grades of Students			
Levels	Grade 7	Grade 8	Grade 9	Grade 10
Prestructural	1.07	3.03	4.17	0
Unistructural	9.35	14.72	10.84	5.19
Multistructural	86.42	79.22	51.67	65.93
Relational	3.19	3.03	33.34	28.89
[No of students	157	77	40	45]

Between Grades 7 and 8, X^2=13.64 at df=3 significant at 0.01 level.

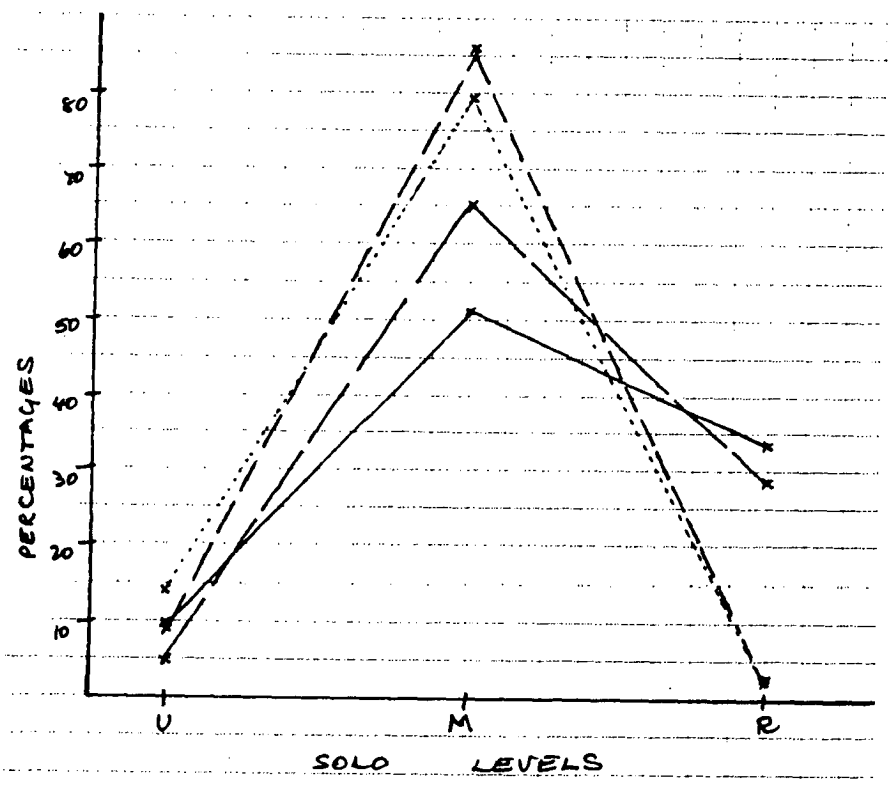
Between Grades 8 and 9, X^2=125.23 at df=3 significant at 0.001 level.

Most of the difference between 8 and 9 was accounted for by Relational responses.

Between Grades 9 and 10, X^2= 13.29 at df=3 significant at 0.01 level.

A statistical comparison of first tests in Grades 8 and 9 using the X^2 Test revealed the difference to be beyond the 0.001 level. There was a fluctuation in the SOLO levels responsible for these differences, which is depicted on Figure 6.04 in diagrammatic form. As can be seen the Unistructural response levels for all grades were similar, as were those overall between the Grade 7 and 8 students. Responses from the Grade 9 and 10 students however were lower at the Multistructural and higher at the Relational levels. As the Grade 9 and 10 students were in elective Music classes this difference could possibly be due to the more intensive musical

Figure 6.04 Comparison of First Test Responses from each Grade



GRADE 7 RESPONSES

GRADE 8 RESPONSES

GRADE 9 RESPONSES

GRADE 10 RESPONSES

training in these courses, to the motivation towards music activities which caused these students to select these courses, or to some other variable. The difference between levels of responses from the two types of Music classes, seen on Figure 6.04, is shown below on Table 6.09 where responses have been combined for Grades 7 and 8, and 9 and 10.

Table 6.09 First Test Responses of Compulsory Grade 7 and 8 Classes and Elective Grade 9 and 10 Classes

	Types of Classes	
Levels	7/8	9/10
Prestructural	1.72	1.69
Unistructural	11.12	7.85
Multistructural	84.05	59.22
Relational	3.15	30.98
[No of students	234	85
Percentage of First Tests	67.15	32.86]

X^2=154.31 at df=3 significant at 0.001 level.

The X^2 Test here, confirmed the observation made from Figure 6.04 that there were significant differences between students in the two types of Music courses, and revealed the main source of the difference as the Relational level. Having established that there were significant differences between grades, the responses of students who were tested twice were examined for further information on the effect on response levels of musical experience and training. Isolating and comparing the data shown on Table 6.10, from the 90 students who were tested twice, revealed that there was a small increase in the number of Relational responses, and a larger increase in the number of Prestructural responses with a corresponding decrease in the number of Unistructural and Multistructural responses.

Table 6.10

Classification of Responses by the 90 Students Tested Twice

Levels	Tests	
	First	Second
Prestructural	0.37	4.45
Unistructural	8.15	7.78
Multistructural	65.93	59.63
Relational	25.56	28.15

X^2=10.56 at df=3 significant only at 0.02 level.

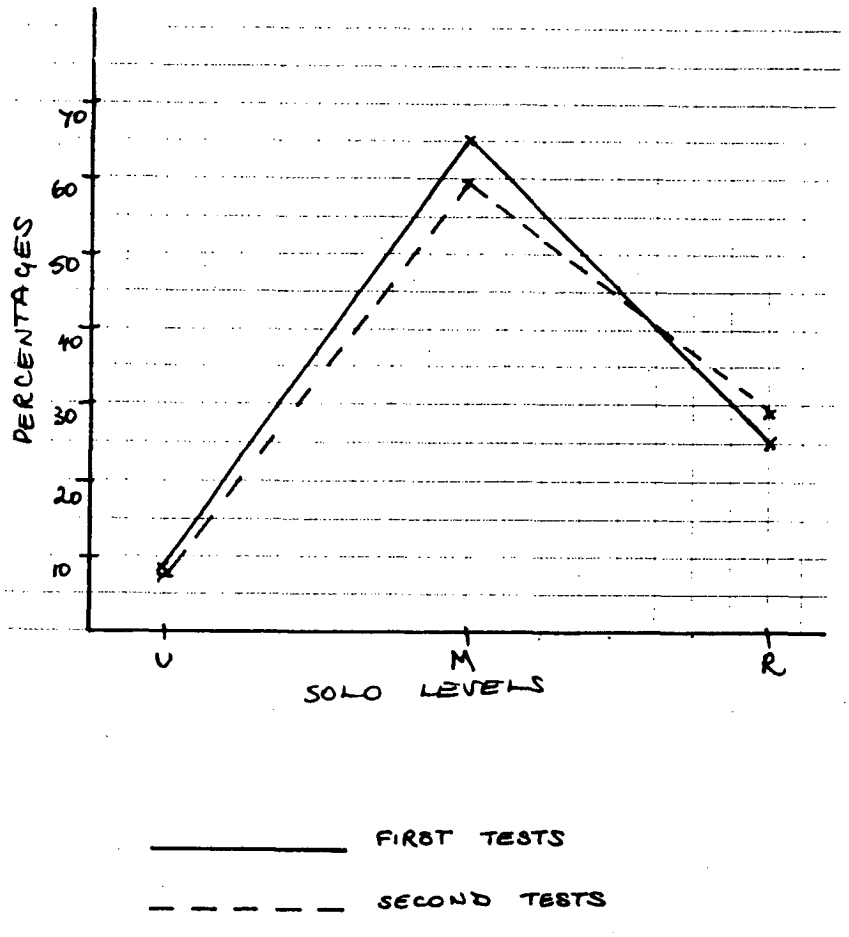
For U,M and R only X^2=0.98 at df=2 not significant.

The significance of these differences between the responses of those students tested twice is not as great as those between the total data for first and second tests, being significant only at the 0.02 level (X^2=9.84) on the X^2 test. The calculated value for X^2 was 10.56, with most of this figure being accounted for by the difference between the Prestructural responses (P^2=9.32). If the Prestructural responses were omitted from the test then these was no significant difference between the groups, thus confirming the observation that the Prestructural responses had been the major source of the observed differences. This can be verified on Figure 6.05 where the patterns for each of these tests can be seen to be very similar.

An examination of the retest data separated into grades was made to determine if there were possible effects due to increased age or musical training. A summary of this examination is presented here, and details can be seen in Appendix K.

Despite there being no Prestructural responses in the first tests and 12 in the second tests of the 48 students retested from Grade 7 to 8, the differences between the tests were statistically not significant. There were only 12 students retested from Grade 8 to Grade 9, so although firm conclusions should not be drawn from their responses, there was a trend in Grade 9 for response levels to be more diverse and the differences were calculated as being significant at the 0.001 level. No significance was found in the differences of responses from the 14 students retested from Grade 9

Figure 6.05 Comparison of Responses from the Students Tested Twice



to Grade 10 or the 13 students tested twice in Grade 10.

If there had been an effect upon SOLO levels due to test familiarity or increased age, then there should have been significant differences between first and second tests in subsequent years. This effect was however only observed between Grades 8 and 9, where it could also have been due to the increase in the amount of musical training from the compulsory 7/8 course to the elective 9/10 course, so the difference here could not be attributed solely to any of these factors. As the numbers of students in each of these comparisons was also small, it would be inadvisable to draw any conclusions at this stage.

There were only 11 students who were tested three times. They were first tested in Grade 8, and then took the test twice in Grade 9 in the Retest Experiment. The responses obtained from these students on each of these tests are shown on Table 6.11. The differences between the levels obtained at each test cannot be tested using the X^2 test due to the scarcity of responses in the lower levels, but the differences can be clearly seen from the data. It seems possible that some of these differences could be due to training, as there was a change from the compulsory Music course in Grade 8 to the elective Music course in Grade 9 (first test to second test), and the SOLO based tuition of the Retest Experiment was interposed between the second and third tests.

Table 6.11

Classification of Responses by the 11 Students Tested Three Times

	Grade in Which Test Was Taken		
	First Test	Second Test	Third Test
Levels	Grade 8	Grade 9	Grade 9
Prestructural	0	3	0
Unistructural	2	5	0
Multistructural	23	13	14
Relational	8	12	19

Retest Experiment

As the 14 students in this experiment were tested three weeks apart, after tuition which focussed upon the musical effects created by combining separate musical elements of music, their results should show not only the effect of training and motivation, but also to some extent be affected by familiarity with the test. The differences between levels on the pretest and posttest can be clearly observed in both the table of collated results, Table 6.12, and in the actual response patterns of each individual student which are shown on Table 6.13.

Table 6.12

Classification of Responses from the Grade 9 Retest Experiment

	Test from which Responses Were Obtained	
Levels	Pretest	Postest
Prestructural	5	0
Unistructural	7	0
Multistructural	17	20
Relational	13	22

On the pretest there were 5 responses at the Prestructural level and 7 at the Unistructural level, whereas on the posttest there were none at these levels. 52% of responses on the posttest were at the Relational level compared to 31% on the pretest. Examination of the individual response patterns on Table 6.13 reveals how these changes occurred. On this table response patterns are grouped according to their pretest results to show the changes in individual student responses. It can be seen that although the responses of all students in the top three lines improved, those whose pretest levels were low (line 1) or whose results had fluctuated (line 2), showed the greatest improvement. The four students (line 4) who were already responding at a high level did not show much improvement on their first levels.

Table 6.13 Individual Response Patterns of the Retest Experiment

Line in Table	Test	Response Patterns			
1	Pretest	UPP	PPU	MUP	MUU
	Posttest	MMM	MMM	MMM	MMM
2	Pretest	RUM	MUR		
	Posttest	RRR	MMM		
3	Pretest	MMM	MMM	MMM	MMM
	Posttest	RRR	MRR	RMR	MMM
4	Pretest	MRR	RRR	RRR	RRR
	Posttest	RRR	RRR	RRR	RRR

After tuition, all students were able to produce responses at least at the Multistructural level. 22 response levels changed from pretest to posttest, and of these only one fell, from Relational to Multistructural (line 2, no 2). All other changes were improvements, mainly from Unistructural to Multistructural or Multistructural to Relational, although 5 response levels improved from Prestructural to Multistructural and one improved from Unistructural to Relational.

At the end of the following year the Schools Board Levels awarded to these students were collected. Table 6.14 below shows the response patterns of these students grouped into their end of course Schools Board Levels. It is interesting to note that although on the Pretest, Unistructural responses had been given by students in each of the groups, and Prestructural responses had been only given by Level 1 students; after tuition, students in both the Level 1 and Level 2 groups produced only Multistructural responses, and most of the Level 3 students produced fully Relational responses sets. A relationship between response levels on the Posttest and Schools Board levels can therefore be clearly seen in this small group of students. The improvement made by the Level 1 students could however be cited to show that the difference in SOLO levels between Level 1 and 2 students might disappear after tuition and strong motivation.

Table 6.14 Individual Response Patterns of Retest Experiment Grouped According to School Board Levels

Level 1	Level 2	Level 3		
UPP	MUU	MUR	MMM	RRR
MMM	MMM	MMM	RRR	RRR
PPU	MMM	RUM	MMM	RRR
MMM	MMM	RRR	RRR	RRR
MUP		MMM	MRR	RRR
MMM		RMR	RRR	RRR

Possible Influences on Test Levels

In an attempt to determine whether there were any strong influences upon the SOLO levels on the test, four factors, Listening Ability, Written Fluency, Motivation and Performance, had been isolated and students who demonstrated outstandingly good or poor skills in these areas had been identified. Although many of the students whose results were isolated for these factors were retested, to eliminate any possible influences due to test familiarity the data used in this section is taken exclusively from their first tests. For comparison purposes data from the students who had not been identified for these factors is included under the heading Average group.

Listening Ability

The groups of students who had been identified for their Listening Ability were compared to each other to test the hypothesis that Music Listening Ability will have no effect upon SOLO levels. Students had been selected by their Music teachers as being either outstandingly strong or weak at identifying musical elements and

talking about the music they heard in class. There were 10 students identified as very weak and 16 identified as being as very strong in this area. As can be seen on Table 6.15 below these two groups produced different patterns of responses. The Low group produced a third of its responses at the Prestructural and Unistructural levels whilst there were no responses at these levels in the High group.

Table 6.15 Classification of First Test Responses (Percentages) in Groups for Listening Ability

	Listening Ability Groupings		
Levels	Low	Average	High
Prestructural	3.3	1.82	0
Unistructural	30	10.13	0
Multistructural	66.6	77.93	75
Relational	0	10.13	25

These differences between the groups are perhaps more clearly shown on the diagram in Figure 6.06. Here the Average group responses can be seen to follow a pattern similar to that observed on Figure 6.01 for all Grade 7 and 8 students. As neither the Low nor the High groups produced responses at all three of the levels shown on the figure, the patterns for each of these groups appear to be onesided and incomplete, and are completely different response patterns from those observed previously.

Despite these obvious differences both Low and High groups had the majority of their responses at the Multistructural level, so it could be that there were strong similarities in the responses of the majority of individuals within the groups. An examination of individual response patterns in the Low and High groups was therefore carried out and is shown on Table 6.16.

Figure 6.06 Comparison of Responses in Groups for Listening Ability

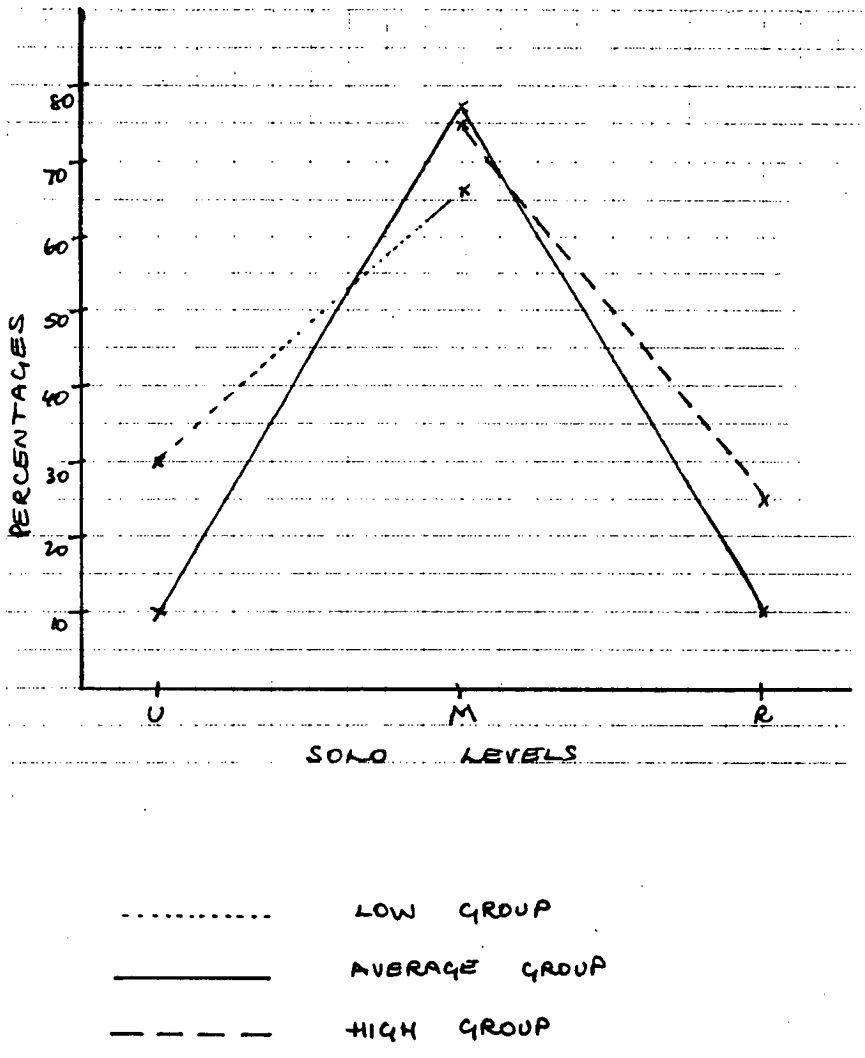


Table 6.16 Individual First Test Response Patterns Grouped for Listening Ability.

(Numbers in brackets indicate number of response sets/grade)

Low Group	High Group
MPM 1 (9)	MMM 9 (8/7,1/ 9)
MUU 1 (8)	MRM 3 (1/7,2/8)
MUM 5 (2/7,1/8,2/9)	RMR 1 (7)
MMU 2 (1/8,1/9)	MRR 1 (7)
MMM 1 (7)	RRR 2 (10)

As can be seen on Table 6.16, the Low and High groups did not contain many students with similar response set patterns. The Low group contained only one student with a fully Multistructural response set, and all other students in this group had at least one Prestructural or Unistructural response. In the High group however there were 9 members with Multistructural response sets, and all other response sets contained at least one Relational response. The two groups obviously contained responses of different qualities. It is also interesting to note that although the best response in the Low group came from a Grade 7 student and the 3 Grade 7s in the High group produced 5 Relational responses between them, the only fully Relational response sets were from Grade 10 students.

Written Fluency

Written Fluency had been defined as the student's ability to express their own thoughts effectively on paper, and it had been thought that this would be a significant influence on student responses. The responses of groups of students who had been identified for their Written Fluency skills were therefore compared to test the hypothesis that Written Fluency would have no effect upon SOLO levels. English teachers had identified 23 students as having very poor skills in this area, and 19 as being outstandingly high achievers in this area.

As can be seen on Table 6.17 below each group produced different patterns of responses. The Low group like the High group had the majority of its responses at the Multistructural level, but

it also had some responses at the Prestructural level, whereas there were none at this level in the High group. There were also large differences between these two groups in the percentage of responses at the Unistructural and Relational levels. It seemed that these two groups had produced responses of different qualities, and this difference was confirmed by the X^2 test to be significant beyond the 0.001 level of probability. The calculated value of X^2 for these groups was 68.21, which exceeds the given value of 18.46 for 4 degrees of freedom at the 0.001 level of probability.

Table 6.17 Classification of First Test Responses (Percentages) in Groups for Written Fluency

	Written Fluency Groupings		
Levels	Low	Average	High
Prestructural	4.34	2.3	0
Unistructural	28.98	8.65	1.75
Multistructural	65.21	73.55	63.15
Relational	1.44	15.52	35.08

With categories P and U combined to remove the empty cell,
X^2=68.21 at df=4 significant at 0.001 level.

The individual response sets within these groups as seen on Table 6.18, seem to fall into into two distinct groups according to their SOLO response levels, with the Low group having nearly all of the Prestructural and Unistructural level responses and the High group having nearly all of the Relational level responses. Each group also seemed to divide into two subgroups. Of the 23 Low group students, 14 or 60.87% of this group, produced at least one Prestructural or Unistructural level response whilst the remaining 9 students, 39.13%, had produced only Multistructural or Relational level responses. In the High group 11 students produced at least one Relational level response, 57.9%, compared to the 8 students, 42.11%, who produced only Multistructural level responses. In each group there was one student whose response pattern did not comform to the general trend. In the Low group this was a Grade 9 student who produced a Relational level response and a Grade 10 student from the High group produced one Unistructural level response.

Figure 6.07 Comparison of Responses in Groups for Written Fluency

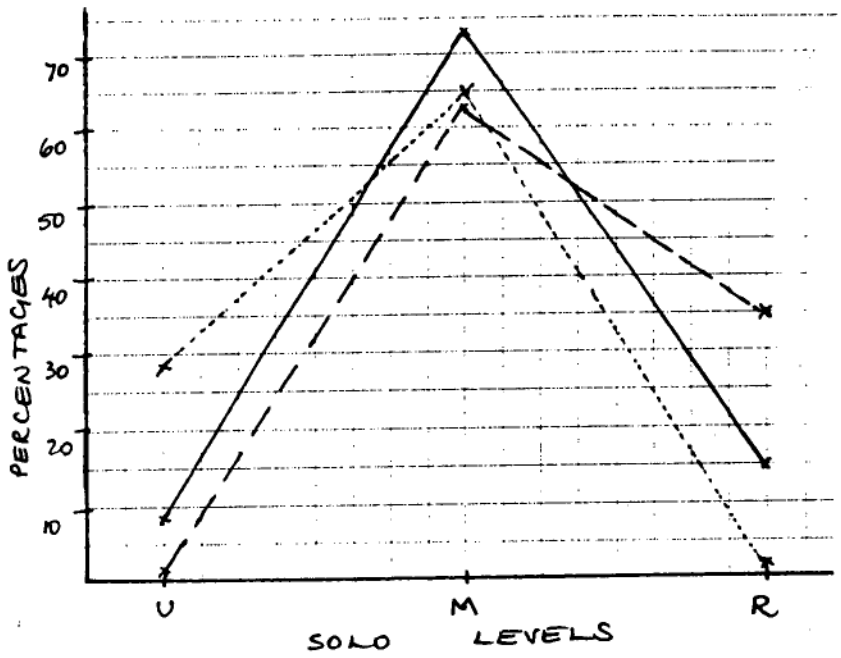


Table 6.18 Individual First Test Response Patterns in Groups for Written Fluency

(Numbers in brackets indicate number/grade)

Low Group	High Group
UPU 1 (9)	RUM 1 (10)
UUU 1 (9)	MMM 8 (5/7,1/8,1/9,1/10)
UUM 1 (7)	RMM 1 (7)
MUU 1 (7)	MRR 5 (3/7,2/8)
MPU 1 (9)	RRR 4 (1/9,3/10)
MPM 1 (9)	
MUM 7 (6/7,1/8)	
MMU 1 (9)	
MMM 8 (5/7,2/8,1/9)	
MMR 1 (9)	

Figure 6.07 gives an even clearer picture of the patterns of responses between the three groups. Here it can be seen that although all three groups had similar numbers of responses at the Multistructural level, there was a tendency for more Unistructural responses from the Low group and more Relational responses from the High group. Although there were differences between the groups there were also strong similarities, and as the individual response patterns on Table 6.18 have shown, individual students did not necessarily follow these trends.

Motivation in Music Lessons

It had been thought that Motivation and attitude towards Music as a school subject would affect response levels and the trends in the response levels of those students identified for this factor seemed to support this theory. 28 students were identified as having poor Motivation in Music lessons and 31 identified as being outstandingly well motivated, and as can be seen on Table 6.19 these two groups produced different patterns of response levels, which when compared using the X^2 test were found to be different beyond the 0.001 level of significance.

Table 6.19 Classification of First Test Responses (Percentages) in Groups for Motivation

	Motivation Groupings		
Levels	Low	Average	High
Prestructural	3.5	2.31	1.07
Unistructural	23.8	8.69	2.15
Multistructural	71.42	73.41	64.51
Relational	1.19	15.33	32.25

X^2=53.63 at df=6 significant at 0.001 level

Once again the percentage of Multistructural responses was similar in each of the three groups, and the largest differences between the Low and High groups appeared in the Relational (31%) and Unistructural (22%) levels. There does seem therefore to be some evidence here to disprove the null hypothesis that Motivation in Music Lessons will have no effect upon SOLO levels.

Musical Performance Ability

The groups of students who had been identified for their Music Performance skills were compared to each other to test the hypothesis that Music Performance Ability will have no effect upon SOLO levels. Students had been selected by their Music teachers as being either outstandingly strong or weak in Musical Performance. There were 20 students identified as very weak and 24 identified as being as very strong for this skill. When compared to each other on Table 6.20 the SOLO response levels of students in these groups were found to be significantly different, a fact confirmed at the 0.001 level by use of the X^2 Test.

Table 6.20 Classification of First Test Responses (Percentages) in Groups for Performance Ability

	Performance Ability Groupings		
Levels	Low	Average	High
Prestructural	3.33	2.4	0
Unistructural	33.33	8.43	5.55
Multistructural	63.33	74.21	55.55
Relational	0	14.99	38.88

With categories PU and MR combined to remove empty cells,
X^2=43.27 at df=2 significant at 0.001 level.

From the table clear differences between the three groups can be observed. The Low group had the highest percentage of Prestructural responses, whilst the High group had none at this level. The Low group also had the largest proportion of Unistructural responses amongst the three groups. Nearly two thirds of responses in the Low group were Multistructural and this also accounted for the majority of responses in the other groups. However whereas the Low group had no Relational responses, this level accounted for 39% of all High group responses. An examination of the individual response patterns shown on Table 6.21 also shows these to be two separate groups.

Table 6.21 Individual First Test Response Patterns in Groups for Performance Ability

(Numbers in brackets indicate number/grade)

Low Group	High Group
UUU 1 (9)	MUM 3 (2/8,1/10)
UPU 1 (9)	RUM 1 (10)
UUM 1 (7)	MMM 7 (6/7,1/8)
MUU 2 (1/7,1/8)	MRM 4 (2/7,2/8)
MPM 1 (9)	MRR 1 (7)RMM 1 (7)
MUM 7 (5/7,2/8)	RMR 1 (7)
MMU 2 (1/8,1/9)	RRR 6 (3/9,3/10)
MMM 5 (4/7,1/8)	

There were however similarities between the two groups on Table 6.21, as MMM and MUM response patterns were found in both groups.

Relational responses were found in the High group but not in the Low group, yet 33 RRR response sets were also found in the Average group; and although most of the Low group had at least one Unistructural response, this level was also found in both the Average and High groups. It seemed that although all groups had the majority of their responses at the Multistructural level, there was a trend towards a greater proportion of Relational levels in the High Musical Performance group, which might indicate a connection between high achievement in this area and SOLO levels for Music Listening.

Relationship between Influencing Factors and SOLO Levels

As the comparison of High and Low groups for each factor had revealed general differences in the SOLO levels achieved by each group, though not necessarily by individuals within each group, it was considered necessary to examine the characteristics of students who had achieved both high and low SOLO levels. As the majority of responses were at the Multistructural level, and most students who produced Prestructural, Unistructural and Relational level responses also produced responses at this level, it was decided to focus on students who produced responses at these less common levels. To avoid duplication and present as accurate a picture as possible, only responses from first tests were considered, and details of this analysis are in Appendix K.

A comparison was made with the total number of students identified for each of these factors to assist in assessing the importance of each factor in determining a Prestructural response. In each case the percentage of students with Prestructural responses was small in comparison to the numbers of students identified for these factors but producing higher level responses. Evidence of strong influences on Prestructural responses seemed to be lacking.

Of those 86 students who produced Unistructural responses on their first tests, 56 were not identified for any factor, whereas 30 were identified for at least one of the possible influencing factors and 25 students were identified in Low groups and 5 in High groups.

The majority of Low group identifications for all four possible influencing factors came from students who had produced Unistructural responses, which seemed to indicate some connection. For instance 8 students with Unistructural responses identified as having exceptionally poor Listening Ability accounted for 80% of those in the Low Listening Ability group, but this was still only a small number of students, and was far outnumbered by students with Unistructural responses who had not been identified for this factor. The situation with the other three factors was even more confused as Unistructural responses were produced by both students identified as having both outstandingly weak and outstandingly strong skills for these factors.

The characteristics of students who had produced Relational responses were examined in the expectation that they would be found amongst the High groups for the possible influencing factors. Although 20 of these students were identified in the High groups, 1 student was identified in two Low groups and there were still 34 students with Relational level responses who had not been identified for any factor. It would therefore seem to be imprudent to attribute the ability to produce a Relational level response to these factors. Amongst these 55 students, 19 produced fully Relational response sets (RRR) from their first test. Of these 19 RRR students, 10 had not been identified for any possible influencing factor, and 9 students were identified 18 times. Although these students were undoubtedly talented they made up a small percentage of the total number of students isolated for each of the factors. For instance 75% of the students identified for outstandingly good skills in musical performance were not amongst these consistently high SOLO level performers. It does not therefore seem feasible to attribute high SOLO levels solely to any of these factors.

SOLO levels compared to Schools Board Levels for Music

As all the 85 Grade 9 and 10 students tested were in elective Music classes following the two-year Schools Board course with three syllabus levels, it was possible to compare SOLO levels from these students with their final Schools Board levels in order to test hypotheses 5.1 and 5.2 concerning the possible relationship of SOLO levels to Schools Board Levels. These students had completed 138 actual tests and produced 414 separate responses, but to avoid possible effects due to test familiarity only their first tests were considered for this comparison. It should be remembered throughout this section that the Schools Board levels were based on the students final level of achievement in all three areas of the Music course, Performing, Composing and Listening, not solely on Listening as were the SOLO levels.

As can be seen from Table 6.22, although the majority of students working at Levels 1 (52.78%) and 2 (76.52) produced responses classified at the Multistructural level, Level 3 students produced the majority of their responses (62.07%) at the Relational level. This observed difference between the students at each of the Schools Board levels was confirmed as significant at the 0.001 level of probability by use of the X^2 Test.

Table 6.22 Classification of First Test Responses from all Grade 9 and 10 Students as Percentages Grouped According to their Final Schools Board Award Levels

	Schools Board Levels		
Levels	Level 1	Level 2	Level 3
Prestructural	8.34	1.52	0
Unistructural	27.78	6.06	2.30
Multistructural	52.78	76.52	35.64
Relational	11.12	15.91	62.07
No of students	12	44	29

With P and U categories combined to avoid the empty cell,

X^2=88.6 at df=4 significant at 0.001 level.

These differences can be seen more clearly on Figure 6.08. The Level Two pattern seems to almost duplicate the overall pattern of responses in this study, whilst the Level One pattern shows more Unistructural and fewer Multistructural and Relational responses than this pattern. The most interesting feature of this diagram however is the pattern caused by the Level Three responses. As the majority of responses had so far been at the Multistructural responses changes in the diagram had been in the degree of the central angle, the time however the majority of responses from the Level Three students were at the Relational level, so the pattern caused by the responses of these students was completely different to any other seen so far, seeming to form a line. This could be taken to indicate a strong direct connection between the SOLO levels and the Schools Board levels.

Two other comparisons of SOLO levels with Schools Board levels were carried out with smaller groups of students during the Main Study, and the results of these supported the main trend seen on Figure 6.07. Data from these smaller groups was included in the overall data presented on Table 6.22.

As has already been seen on Table 6.14, the Schools Board levels of the students in the Retest Experiment class followed this trend with all Relational level responses being given by Level 3 students, and 7 of the 9 students at this level producing fully Relational response sets after tuition.

A comparison of the Schools Board levels and SOLO levels for two parallel Grade 10 Music classes was carried out as part of a Moderation procedure to check that Schools Board levels had been allocated fairly, and this is been documented on Table 6.23 From this table it can be seen that although the numbers of students involved was small, the trends shown on the previous tables were followed with most of the Relational level reponses being produced by the Level Three students.

Figure 6.08 Comparison of First Test Responses from Grade 9 and 10 Students According to Final Schools Board Levels

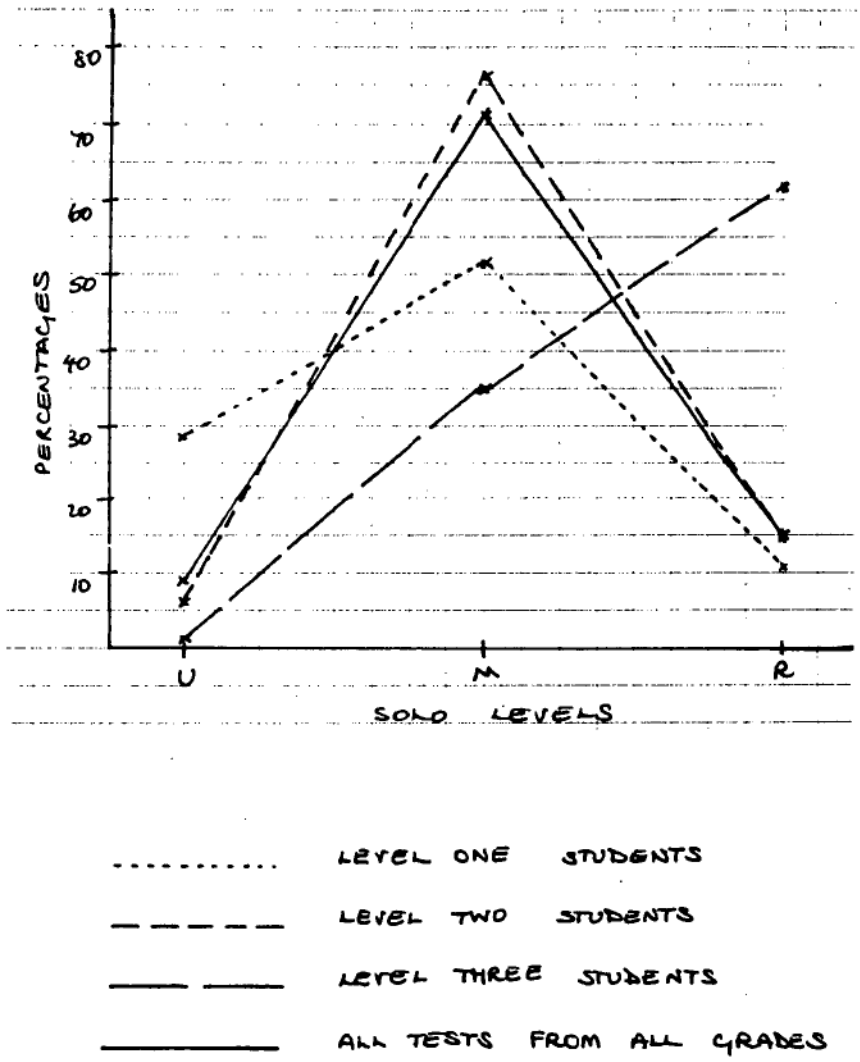


Table 6.23 Individual Response Patterns in Moderation Experiment Grouped According to Schools Board Levels

Level 1	Level 2	Level 3
MUM 1	MUM 2	MRR 4
MMM 2	MMM 10	RRR 7
	RMM 1	
	RRR 1	

All the level 3 students had at least 2 Relational responses, whereas only two level 2 students had response sets which included a Relational response, and the majority of level 2 response sets were based around the Multistructural level. It was not however impossible for a student with a Level 2 award to demonstrate listening skills at the Relational level, and one Level 2 student produced a fully Relational response set. In the analysis of responses on Table 6.24 the numbers of actual responses have been given in preference to percentages which as only 28 students were involved could give a misleading impression.

Table 6.24 Classification of Responses in Moderation Experiment Grouped According to Final Schools Board Levels

	Schools Board Levels			
Levels	Level 1	Level 2	Level 3	
Prestructural	0	0	0	
Unistructural	1	2	0	
Multistructural	8	36	4	
Relational	0	4	29	
[Total No Responses	9	42	33	
Total No Students	3	14	11	28
Total % Responses	11%	50%	39%]	

[Statistical verification of differences not possible due to the empty cells.]

The data from these three individual classes (1 Grade 9 Retest, and 2 Grade 10 Moderation) supported the overall picture obtained from the total data for students in Grade 9 and 10 elective Music classes. As the level 3 syllabus had objectives in each of the 3 Music course areas, Performance, Composition and Listening, which focussed on the interaction between musical concepts and could be

held to require thought at the Relational level, this trend towards a convergent relationship between Relational level responses and students assessed at Level 3 was the one that should have been noted if the SOLO test was an accurate measure of thought at this level.

Musical Elements used in Responses

As it had been observed in the Pilot Study that musical elements in the second group on the Elements Checklist seemed to be the focal point of Relational responses, hypotheses about this and the possibility that teachers might be able to judge the quality of a response by observation of the type of elements selected had been formulated. An analysis of the types of elements mentioned in the responses which used musical elements was carried out to test these ideas, and the results can be seen on Table 6.25.

From this table it seemed that responses at the Unistructural level had some obvious characteristics. They mostly identified the basic elements of music such as instruments, speed, volume and mood and imagery.

The comparison of elements mentioned in the Multistructural and Relational responses was more complex. The observation made in the Pilot Study that Relational responses mentioned more Group Two and Three elements seemed to have been borne out. A statistical examination of the Group Two and Three elements mentioned in Multistructural and Relational responses made with the X^2 Test to investigate the null hypothesis that there was no difference between the elements mentioned in these responses, rejected this hypothesis at the 0.02 level of probability. This rejection meant that the probability of the differences occurring by chance was 2 in 100, it did not give any indication that an individual student would necessarily follow this trend. Although, with the exception of performance practice, every element in each group was mentioned more often in Relational responses, they were all also mentioned in some Multistructural responses.

Table 6.25 Comparison of the Musical Elements Mentioned in Responses

[All figures are percentages of responses at that level mentioning the element listed.]

Elements	Response Levels		
Group One Elements	Unistructural	Multistructural	Relational
Instrument/Voice ID	19.86	58.86	91.86
" " Group	0	22.49	51.75
Speed	15.45	65.55	68.03
Volume	15.45	60.77	88.96
Melodic elements	8.83	44.99	76.17
Rhythm and Beat	2.21	18.66	23.26
Mood	19.86	57.90	75.59
Imagery	14.71	56.46	58.14
Timbre *	0	10.53	41.28
Formal elements *	2.21	56.4	85.47
Group Two Elements			
	Unistructural	Multistructural	Relational
Orchestration	0	32.06	72.10
Formal Structure	0	32.06	72.10
Texture	0	26.80	79.66
Tonality	0	3.83	13.38
Style	1.47	16.27	18.61
Group Three Elements			
Performance Practice	0	0	1.17
Composition Principles	0	1.48	3.49

[Group 2 and 3, M and R responses only : With last two categories combined to remove the empty cell, X^2=15.06 at df=5 significant at 0.02 level.]

*Timbre, the description of specific sound, is included here in Group One, despite being listed for convenience in Group Two on the checklist. Orchestration, the explanation of how specific sounds are used in combination with others to create desired effects, is an element requiring combination of other elements and therefore remains in group 2.

*Formal elements, is the category in which comments such as "*at the start*", "later on" *and* "*it finished with...*" which identified a passage of music in time were placed. Formal Structure is the category where the relationship of passages or sections in the music to each other was explained. The recognition and explanation of a formal structure is therefore a skill which requires the understanding of the relationships between musical ideas, whereas the comments noted under formal elements belong in Group One. Both timbre and formal elements were listed in Group Two on the checklist to enable discrimination from the closely related relational categories to be made swiftly.

*It should also be noted that the figures for Imagery in the M and R columns included piece no 2, Varèse, which required an image to be mentioned. This was the only question which focussed students' attention on a particular element.

The Group Three elements, which it had been thought might form the basis of Extended Abstract thoughts about Music, were used so seldom by students that it is impossible to draw any conclusions about their use. This study did not contain any Extended Abstract responses, so it is also not possible to refute the hypothesis that these musical elements could be the basis of responses at this level.

The same Group One elements were used in all responses, though they were all used more often in Multistructural responses than in Unistructural ones, and more often again in Relational responses. The same trend between the Multistructural and Relational responses can also be seen on Table 6.25 with the Group Two and Three elements. As the same elements were used in responses at different levels it does not therefore seem possible to use elements as the sole basis for discrimination between responses as has been done on point scoring marking schemes, but there is no doubt that the more complex Relational responses did use more of the more complex Group Two and Three elements. It seemed that a relationship between the groups of elements and SOLO levels could be seen, disproving the hypothesis that musical elements could not be classified into SOLO levels. As higher level responses seemed to use elements from all group whereas lower level responses had tended to use more elements from Group One than from Groups Two and Three, there seemed to be no evidence to refute the hypothesis that SOLO levels would have no effect upon the groups from which the elements within it were taken. As the SOLO levels are based on the increasing complexity of the operational characteristics of responses, the use of the more complex elements which integrate simpler elements can be expected at this level, and could perhaps be used by teachers as an indicator that the student might be operating at the Relational level.

Chapter Seven

An Examination of the SOLO Music Listening Evaluation

The Music Listening evaluation technique used in this study consisted of three pieces of music which were played to students three times. An open-ended question was set for each piece and answered in writing by each student, and these responses were then classified into SOLO levels using the Guidelines for SOLO Assessment of music responses as set out in Chapter 4 (Figure 4.02). In this chapter this technique will be examined to determine the reliability, validity and feasibility (useability) of its use in the Music classroom.

1 : Reliability

Reliability, the degree of consistency between two measures of the same topic, can be tested in various ways: through studying the equivalence of scores from parallel tests; by examining internal consistency; by comparing the scores allocated by different judges or the same judge upon different occasions; and by looking for stability of scores from tests of the same students made at different times. As there was no suitable Music Listening test available, the equivalence of a parallel test cannot be discussed. The nature and length of the SOLO test also rules out any measures of internal consistency, so discussion of reliability must focus on intra- and inter-judge reliability and the stability of scores on the different pieces and the retests.

Intra-judge Reliability

Although decisions about the format of the testing technique had been made on the basis of a year of classroom observation and trials, which had generated many responses, the only responses included in the data for this study were those obtained from the formal test situations outlined in chapters 4 and 5. These tests generated 1305 individual responses (1260 Main Study + 45 Pilot Study) which were all were classified according to the SOLO levels by the investigator immediately after the test sessions in 1987, 1988 and 1989. All 1305 separate responses were reassessed again at the end of the Main Study in March 1990. On both occasions all responses received the same SOLO classification from the investigator, showing that the SOLO assessment guidelines were capable of being interpreted in the same way by the same person despite an intervening gap of several years, and establishing intra-judge reliability.

Inter-judge Reliability

Responses were also classified into the same SOLO levels by nine experienced secondary music teachers, who were involved as judges of SOLO levels on 3 separate occasions. Upon each occasion responses were given the same SOLO levels. The first of these judging sessions occurred during the Pilot Study, when two teachers first assessed responses according to their usual methods and were then given the SOLO Guidelines and asked to classify the same responses into SOLO levels. They were unanimous about the SOLO classifications, as were the group of six teachers who repeated this procedure in the middle of the second year of the Main Study. The test was also used as part of a Moderation procedure between two parallel Grade 10 classes with four teachers, three of whom had already been involved once before. Upon this occasion they classified responses according to both SOLO and conventional marking systems. SOLO levels were once again consistent between judges.

It does seem therefore that, not only is it possible to use SOLO to classify students responses to music, but that one judge will make the same classifications on different occasions, and that different judges will classify responses into the same levels. The conclusion which can be drawn from this is that the standard of evaluation of Music Listening responses which was set by the use of the SOLO Guidelines was a constant and reliable one.

Stability of SOLO levels across response sets and retests

If the test is reliable and the standard set is constant, then the levels achieved by individual students should show a tendency to remain constant over retests and over different stimulii. As SOLO levels are not intended to classify a student, but to identify the standard of a response to a particular problem at a specific moment in time, it might be expected that response levels would differ according to the piece of music, and this assumption was part of the reason behind the inclusion of three pieces of music in the test. However a trend towards the production of responses at the same SOLO level was observed, with 59% of response sets having all three responses classified at the same SOLO level, and 97.4% having two or more responses at the same SOLO level. It does therefore seem that there was a strong trend for SOLO levels to demonstrate reliability through stability of levels within responses sets.

It was also expected that responses from students tested on consecutive occasions would not be the same, and that there would be some increase in levels possibly due to test familiarity, greater experience as musical tuition continued or an increase in skills due to age. The results from the first and second tests however, which were depicted in Figure 6.05, seemed to be very similar, with the significant statistical difference between these tests being attributed to responses at the Prestructural level.

An examination of response levels between retests in each grade revealed interesting trends, although as the numbers in each of these groups were small, these observed trends should not however be given as much credence as those from the overall retest data. No significant difference was found between retests from Grades 7 to 8, or from Grades 9 to 10, or from retesting carried out during Grade 10, and this could be held to indicate that response levels tended to remained stable across retests. Conflicting data, however, emerged from an examination of the response levels of the students tested in Grade 8 and retested in Grade 9. The observed differences between the response levels in these retests, whilst collected from only 12 students, were found to be significant beyond the 0.001 level. As the Grade 8 students had been tested in a general Music course, whereas the Grade 9 students were in a more intensive elective Music course, the possibility existed that the differences in response levels had been influenced by the change in the type of Music course. Differences between groups following the same courses were very small, and although the group tested from Grade 8 to 9 was small, the results here could be taken to demonstrate both the stability of the test under equivalent conditions and the possible sensitivity of the test to changes in Music tuition.

All this is evidence in support of the reliability of the SOLO assessment technique. Not only could Music Listening responses be assessed using SOLO levels, but different judges awarded the same SOLO levels, and ranked responses in the same order using their normal assessment methods and SOLO levels. Students also showed a tendency to give several responses at the same level during a test session, and, in the absence of changes in Music tuition, to tend to remain constant over retests. In the absence therefore of any conclusive evidence to the contrary, the hypotheses that SOLO levels remain constant and that this evaluation technique is a reliable device, seem to hold true.

Significant effects on response levels were found, though only in small groups, which could possibily be atributed to changes in the type of Music course and tuition. As this SOLO evaluation

technique is intended to be used in the normal school situation, it should be sensitive to both instruction and motivation, and this demonstrated sensitivity should therefore be seen as an important argument in support of the validity of its use in the classroom situation.

2 : Validity

As was shown in Chapter Six, by comparison of responses from students in both compulsory and more intensive elective Music classes and from the Retest Experiment group, SOLO levels on this test could be sensitive to and improve with an increase in both general musical training and in specific Music Listening training designed to emphasise the interrelationships between various aspects of Music. This could be taken as an indication that this SOLO test has achieved its aim, is a measure of what is being taught in Music courses, and therefore has validity for use in secondary Music classrooms.

Face Validity

In the initial stages of test construction, during the Pilot Study, practising music teachers were consulted, asked to comment and offer suggestions about the test format, the pieces of music and the questions which were to be used to obtain the responses. All the teachers consulted considered the materials used and the method of obtaining responses to be the equivalent of listening tests which they had used to evaluate students of this age and experience. Despite the strong influence that language skills might play due to the written format of the test, which was seen by these teachers as an unavoidable problem with any school based test, they also considered that the SOLO test was a suitable method of assessing the Listening component of the current Schools Board course.

Construct Validity

It is also important to re-examine in the light of the collected data, the hypotheses linking the SOLO Taxonomy and this test, and to establish its construct validity. As was explained in Chapter One, the study began by examining the qualities that characterise a good listener. An experienced listener, such as a professional musician, music critic or musicologist will have a wealth of previous aural experiences with which to compare the current one, and will focus on more complex musical structures such as the principles of composition, which imply understanding of simpler musical concepts such as tune recognition. Abstract theoretical speculation, as seen in journals such as *Sounds Australian* is clear evidence of functioning at the Extended Abstract level, whilst listeners who can bring together awareness of the separate elements into the recognition and appreciation of complex concepts such as orchestration, form and style are demonstrating thought at the Relational level, and listeners who do not see further than the simpler separate musical elements are operating at the Unistructural or Multistructural level.

In the data from this test not only were these distinctions between responses clearly observed by all the judges, but the analysis of musical elements used in the responses (Table 6.25) showed the parallels observed between the increasing complexity of cognitive structures in the types of musical elements and the increasing complexity of operations in the SOLO levels. This had also been observed by Biggs and Collis (1982) and noted in their work on the SOLO Taxonomy. There does therefore seem to be a relationship between the theoretical work on SOLO and the results obtained from this study, which would support the validity of the SOLO Listening test.

Content Validity

The next point to consider is whether this test really does require a student to demonstrate the skills set out by the Schools

Board course statements. The directive to teachers cited in Chapter One was :

> *"Teachers will look for evidence that students can : identify, describe and compare features of music, and show an informed awareness of design and techniques used in creating an overall effect."*

The instruction given to students was to *describe* the music, and as was seen on Table 6.25 most of them did this by *identifying* and *describing* several features of the music in their Multistructural responses, thus fulfilling two of the assessment points noted above. Those students who noticed Group Two musical elements had undoubtedly recognised them by *comparison* with music previously heard, but only those students with responses classified at the Relational level had succeeded in explaining their comparisons or how the overall effects were created. Evidence of *awareness of design and techniques used in creating an overall effect,* which was the aim of the question about the piece by Varèse, were shown to some extent in Multistructural responses but were only explained satisfactorily in the Relational level responses. It did seem that the SOLO response classifications were revealing the extent to which the student was able carry out these required tasks.

Responses to the task given could therefore be shown to fulfill the assessment directive given to teachers, but were the materials used to stimulate these responses typical of those used in the Music courses? In order to stimulate the widest and highest level of responses possible from students the pieces selected for the Main Study were limited to orchestral music, and were not truly representative of all the music used in these courses, but were nevertheless still considered by the music teachers involved to be similar to the listening repertoire used in other schools. As reported in Chapter Four however, a wide variety of music similar to that used in other High School Music courses was used in the Pilot study to stimulate responses and it was found that responses could be classified into SOLO levels, so the limited nature of the music used in the Main Study should not be taken as an indication that this is the only type of music suitable for use with this test.

If the task and materials used were similar to those used in other Grade 9 and 10 Music courses, and the responses produced gave evidence of the points required by the Schools Board, then the test could be held to fulfill its aims, unless there was evidence that responses had been influenced by factors unrelated to the Music Listening skill being evaluated. In an attempt to establish whether responses on the test were representative of student responses to Music Listening or were influenced by extraneous variables, some possible influencing factors had been isolated.

As was discussed in Chapter Two there are innumerable possible influences upon every listening experience and therefore upon responses to this test. The factors isolated here were those which like Written Fluency or Motivation might distort the level of student responses, showing that the test was not in fact a good measure of Music Listening; or which might, by demonstrating a congruent relationship with a listening related area, such as Performance Ability, Music Listening Ability, or an increase in Music Listening tuition, support the validity of the SOLO test.

As the SOLO test required a verbal response it was subject to the influence of the student's language ability, and as responses were written down they were also probably influenced by the students' ability to write quickly and clearly, to be able to read their own writing, to express their thoughts on paper and to review what they had written effectively. The expectation behind the isolation of this factor therefore had been that students with good language skills would produce higher SOLO levels.

The majority of responses from both groups identified for Written Fluency however were at the Multistructural level, though there were more Unistructural responses in the Low group and a large proportion of Relational responses in the High group. Despite this a greater percentage of Relational responses were produced by students in Grades 9 and 10, showing that Music tuition was probably a stronger influence upon response levels than this factor. The data from this study, as reported in Chapter Six, seems to suggest therefore that although Written Fluency might be a contributing

factor towards response levels, it was not necessarily a limiting factor on the achievement of Multistructural or Relational level responses.

This hypothesis that Written Fluency was not a limiting factor upon the SOLO levels in this test was supported by the responses of one student, a Laotian refugee, who at the time of the first test had very poor English language skills. This student consistently produced high level responses, which despite errors in grammatical expression and spelling, gave evidence of intensive listening and showed an awareness of the intrinsic relationships between separate musical elements. This particular student was tested three times and produced 1 Multistructural and 8 Relational level responses, a performance only equalled by one of the other 11 students who were tested three times. Problems with language and written expression had not, in this case, restricted the student's ability to listen in depth and explain what had been heard. During the period of the study this student also learned to play several instruments and was obviously very keen on Music and therefore motivated to work hard at this subject. It could be that this student's attitude and consequent motivation had enabled the language barrier to be overcome. The role played by motivation in influencing SOLO levels was also examined for all students.

As it had been noticed by the collaborating teachers in the Pilot Study that the students isolated for poor Written Fluency skills were also poorly motivated in Music lessons, this had been isolated as a factor in the Main Study. When examined the response levels of the students isolated for Motivation towards Music Lessons showed that although there was a significant difference in response levels between the groups, each group produced responses at each of the four SOLO levels. As usual the majority of responses in each group were at the Multistructural level, and the difference was in the percentage of responses at each level, not the types of responses found in each group as in the groups previously isolated. which would seem to indicate that high motivation may not be an essential factor in the attainment of high SOLO levels. Once again the percentage of Relational responses achieved by the High group

was lower than that from the students in each of the elective Music grades (9 and 10), showing once again that musical training could possibly be a stronger influence than motivation. It cannot therefore be asserted that SOLO levels are dependent upon student motivation, but it cannot be denied that motivation may be a contributing factor in the production of higher and lower level responses.

Intensive listening skills are essential for a good performance of a piece of music, so there was also an expectation that Musical Performance Ability would be an influence on test results. Musical Performance requires a high level of cognitive ability in the coordination and integration of visual skills in the deciphering of music notation for sight reading; aural skills in the determination of the correctness of pitch, rhythm, tone and intonation; and kinetic skills involved in the physical control of the musical instrument. The identification of an outstanding student in this area might, therefore, however be indicative of a group of complex factors rather than one aspect of musical ability. Students involved in musical performance are already exhibiting cognitive relational skills linked to a practical endpoint, and those identified in this study in the High group for this factor were those starting to include personal interpretation in their performances. They were beginning to adapt the initial stimulus of the printed or heard music into a personal statement by minute alterations in aspects such as tempo, phrasing, articulation. They were therefore starting to show signs of Extended Abstract thought in their performances, so there was an expectation that this group would produce the highest SOLO levels.

The High Music Performance group did in fact not only produce higher levels of responses than the Low level group, but also produced a percentage of Relational level responses which was higher than that for the overall sample, but still not higher than that produced by all the Grade 10s tested. This could indicate that although Music Performance Ability was strongly related to SOLO levels on this test, the SOLO levels had a stronger relationship with the amount of musical training. Support for the congruent

relationship between Music Performance and the SOLO levels could also be seen in the fact that the High group produced no Prestructural responses whereas the Low group produced no Relational responses.

If the technique did evaluate Music Listening skills then there should also be a convergent relationship between the SOLO levels obtained on this test and other classroom assessments of Music Listening. The possibility of this relationship existing between SOLO levels and conventional music assessments was investigated in two ways : by examining the SOLO levels of students identified for their Music Listening abilities; and by comparing the end-of-course Schools Board levels obtained by the Grade 9 and 10 students with their SOLO levels.

When the SOLO levels of the students who had been identified for their ability to discuss pieces of music in class were examined, there were obvious differences between the Low and the High groups. Although the majority of responses from these groups, as from all others, were at the Multistructural level, the Average group produced responses at all levels, and as was seen on Figure 6.06, the High and Low groups did not follow this pattern but had the rest of their response at the Relational or Unistructural/Prestructural levels only. This would seem to indicate that there was a strong relationship between the two measures of Music Listening in this small sample of students, indicating content validity.

There is also a possibility that SOLO levels on this test were sensitive to the amount of music tuition that the students had received. This seems to be indicated by the retest data where those students who were first tested in Grade 8 during the compulsory general Music course and then retested in Grade 9 in the elective Music course, were the only ones to show a significant improvement. As these were the same students the influence upon their levels of many influences, such as Written Fluency, Performance Ability and Motivation, can be discounted and the significant difference in levels is more likely to be attributable to the change in the type

of music tuition received.

Specific tuition in Music Listening may also have been an influencing factor in the improvements in SOLO levels from Pretest to Postest which were observed in the Retest Experiment. Here 10 of the 14 students improved their response levels by at least one level, and in several cases by two levels; and the only students who whose levels did not improve were the four students who achieved fully Multistructural or Relational response sets on the Pretest. Once again the sample of students was small, but a pronounced trend towards an improvement could be seen, which seemed to indicate that it was possible to influence the SOLO levels which students produced in response to a piece of music. Whether the improvements were due to the influence of the specific tuition, the practical work-in-progress listening being carried out at the same time, an increase in motivation or some other variable could not be determined from this brief experiment. It can however be stated that the test was able to measure the change in responses to music which took place from the pretest to the posttest, and as the standard of responses after tuition is what music teachers need to be able to measure, this can be taken as another indication of content validity.

As the Schools Board Level Three syllabus and assessment guidelines stressed skills and concepts which involved the interaction of simpler concepts, it was expected that students being assessed at Level Three would produce most of the Relational level responses amongst Grade 9 and 10 students. In the overall data 16% of responses were at the Relational level, and 45% of all Grade 10 students gave responses at this level, which seemed to provide further evidence for the relationship between the SOLO levels and amount of musical training. On their first tests 62% of the Level Three students achieved this level as compared to 16% of Level Two students, which gives further support to this theory and also demonstrates the strong relationship between the SOLO levels and Schools Board levels already seen on Figure 6.08. Once again there was a positive correlation between the two measures, SOLO and Schools Board levels, which seemed to indicate that the SOLO levels were reflecting the current official assessment of students.

Despite the strength of the overall data, when individual student results were examined it was found that the general trend was not necessarily followed by individuals as some Level Three students had produced Multistructural response sets, and some Level Two students had produced 2 or more Relational responses. The students had not all followed the trends, so whilst the SOLO levels bore a strong relationship to the School Board assessment pattern, they did not give exactly the same picture of the students' attainment. As the Schools Board assessment had been arrived at by combining results from a two year course in Listening, Performing and Composing, it would be unrealistic to expect it to mirror exactly SOLO levels produced from one test in one area, and the extent to which it seemed to do this is really quite remarkable.

Content validity is concerned with the degree to which the test requires the student to demonstrate the skill being examined, and it has been shown here that the materials and task used in the test provided the opportunity for the student to demonstrate the skills required by the syllabus. It has also been shown that although Written Fluency and Motivation may be contributing factors in the production of Relational responses they are not limiting factors, and are not seen as distorting the results. Evidence of congruent relationships between the SOLO levels and music teacher assessed factors such as Performance Ability and Listening Ability, also seem to indicate that the students tested were demonstrating Listening skills as understood by their teachers. The strong relationship demonstrated between SOLO and Schools Board levels also seems to indicate that both types of classifications were providing the same type of information about students.

Factorial and Predictive Validity

Due to the nature of the test used it is not appropriate to investigate factorial validity. As the aim of the study was to devise a measure of a student's achievement in Music Listening at a

precise moment within the Music course, predictive validity was not considered in the planning of the study. The close correlation between SOLO levels and the final Schools Board levels, which for the Grade 9 students were awarded a year after they took the SOLO test, seems however to indicate that the SOLO test might possibly have some predictive value.

3 : Useability

In this study useability was taken to be the extent to which the SOLO test was suitable for use in Tasmanian secondary school Music courses. The representative nature of the student sample, student and teacher opinions, and the practicality of its use in the classroom were all considered.

Although the students used in this study were selected on the basis of availability, they were a representative sample of students in Music courses in Tasmanian High Schools. The Grade 7 and 8 class groupings had been made at the beginning of the year on the basis of heterogeneity of sex, intellectual ability, attitude and behaviour, and so were representative of the population in these grades. The Grade 9 and 10 students however had all selected Music as an elective subject, and probably therefore were the students with the most motivation and musical talents in these grades. However, as the original aim of the study had been to devise a procedure for use with elective Music students, it could be held that the sample used in the study was in fact representative of the target population.

All students found the test simple to complete, and reported in the informal discussions after each test that they felt comfortable with the procedure, and they were certainly happy with the SOLO based feedback which was provided on their responses. At the end of each year the method was explained and discussed with the Grade 10 students, who were not to be tested again, and they were offered the option of having their work reassessed using a conventional marking scheme. All the Grade 10 students understood the principles involved, thought the SOLO assessment system was fair to them, and

were willing to accept it as a method of assessing their work. The Grade 9 Retest Experiment Group also discussed the principles involved and thought it was a fair method of assessment. As these students had practised with other pieces, had their work marked and discussed their levels thoroughly before the Posttest, they were perhaps in the best position to comment on its use. It seemed that not only was the test procedure acceptable to students, but that they were willing to accept an assessment of their work based on the principles of the SOLO Taxonomy.

The teachers involved with the study thought that once they had become accustomed to using SOLO for assessment, they would find it easy and fast to use, and with one exception, they thought that it would probably provide a more equable assessment than their normal methods. The teachers who had been involved in the Moderation assessment of the parallel Grade 10 classes, also felt that the SOLO levels awarded for this one test had been an accurate reflection of the standard of the students' work throughout the year.

As the technique collected responses in the same way as existing assessment techniques with the entire class involved at the same time, and involved only equipment standard in Music classrooms, the teachers also saw it as being inexpensive, simple to administer and practical for use with their classes. They commented on ways that they could adapt the tasks so as to focus on different aspects of the music, whilst still retaining the same assessment format. They also offered suggestions as to other listening activities in a normal Music lesson, which could be assessed using the same principles. Investigations with other listening activities, which had been carried out during the Pilot Study and detailed in Appendix C, had already shown that SOLO had possibilities for application with many of the procedures commonly used in music classrooms to evaluate listening skills.

These teachers also felt that feedback based upon SOLO, similar to that provided for students in the Main Study would be simple for them to organise, easy for students to understand and would both provide students with a meaningful record of progress and also

create an atmosphere of positive motivation. They also saw another use for the SOLO based feedback, as they thought that it could be used quite legitimately to give adequate grounds for the explanation and justification of awards, and to explain these to parents. The convergent relationship between SOLO levels and Schools Board levels, which was observed both in the Moderation Experiment and in the data from all the Grade 9 and 10 students, was another factor which supported teachers feelings that this test would be very useful with students in elective Music courses.

The brief Retest Experiment, although carried out with only 14 students, seemed to open up another possibility for the use of this evaluation technique, that of an aid to the teaching of active listening. Most student response levels improved from Pretest to Posttest, and all students reported feeling happier about this area of the course after the tuition, so as a motivation exercise it was successful, though the cause of the improvement in response levels could not be determined. Although the cause could not be pinpointed, the results from this experiment showed some trends that could be worth further investigation.

Summary

In this chapter results from the two studies have been cited to demonstrate the reliability, validity and useability of the SOLO Music Listening evaluation technique. Reliability was established through an examination of the judgment of response levels by different judges and the same judge upon different occasions, and through the stability of response levels across pieces and retests. Face validity was established by the opinions of music teachers, and construct and content validity were supported by the nature of the response levels produced by different groups of students. The validity of the application of this test to other students, useability, was established through an examination of the population samples, the opinions of students and music teachers, and the suitability of the test for use in the Music classrooms and courses of Tasmania.

Chapter Eight

Conclusions and Implications

Music Education throughout the world recognises three main areas of activity : Listening, Performance and Composition, and requires that they all be evaluated effectively. As Music is an important part of the social life of most cultures, Listening is not only the skill upon which the other two areas depend, but also the one which everyone continues to use after they have left formal education and in which many people choose to dedicate large amounts of their leisure time. As Listening was not only such an important area, but the one which had been given the least attention in the Tasmanian Schools Board Music Moderation meetings, it was this area upon which this study focussed.

The aim of this study, as stated in Chapter One, was to develop an effective evaluation tool for Grade 9 and 10 Music courses in the area of Music Listening. A test to assess responses to Music was required that could be used in the classroom situation. It would also be desirable if the test was not only simple, quick and easy for teachers to use, but also made sense to students and could be seen by them to make a fair assessment of their work.

A search of the literature dealing with Music Listening showed that although much work had been carried out in this area, most of it had been concerned with the ability to recognise the separate elements of music, had often used artificial aural stimuli, and was unsuitable for regular classroom use. Investigations using extracts

from complete pieces of music had often classified responses by the type of elements used without rating the relative importance of each type of response, but Gardner's (1973) study of paired extracts showed that his classifications seemed to fall into a hierarchical structure linked to age and possibly to musical training, which seemed to show some relationship with Copland's three planes of increasing skill in listening to music.

The structural complexity of student thinking was also examined for a possible basis to an evaluation system, and the SOLO Taxonomy was found not only to have possibilities in this areas, but also to show similarities with some theories of the development of students' thinking about Art and Music (Parsons 1987, Ross 1984, Swanwick and Tillman 1986). SOLO also seemed to provide a theoretical basis for the types of responses observed by Nelson (1984) and Gardner (1973). the classification of aural test items made by Colwell (1987), and for the Music Listening Taxonomy developed by Goolsby (1984). As a brief informal analysis of some student responses to music showed that it seemed to be possible to classify responses into SOLO levels, it was decided to investigate the application of the SOLO Taxonomy to the evaluation of Music Listening in the classroom further by developing a SOLO Music Listening test and using it with secondary school students.

This test was developed through a one year Pilot Study which showed that although SOLO responses obtained from different situations could be evaluated using SOLO, the type of task set had a strong influence upon response levels. Recall questions such as *What?, When?* and *Who?* stimulated Unistructural responses, whilst higher level questions such as *How?* and *Why?* tended to stimulate responses at many levels, gave the student the opportunity to select their own level of response, and provided the teacher with more assessment information than the student's ability to recognise and identify sounds. Trials of the test format and evaluation procedure carried out with the collaboration of experienced secondary music teachers culminated in an end of year test with Grade 10 students. Some possible influences upon responses levels were considered and these were investigated further in the Main Study.

In the Main Study students in Grades 7-10 in two Tasmanian High Schools were tested, with many students being retested. Response levels from the elective Music students in Grades 9 and 10 were compared to their final Schools Board levels and responses from parallel Grade 10 classes were compared as part of an inschool Moderation procedure. A brief experiment involving specific tuition in the principles behind the SOLO Taxonomy was also carried out, and the influence upon SOLO levels of Written Fluency, Motivation, Performance Ability and Listening Ability were also examined.

The test which was developed through this study was based upon the SOLO Taxonomy, and was used to classify student responses to Music into broad categories for evaluation purposes. Despite the presumed influence of other factors such as written fluency upon the responses obtained from students, the use of this technique and its written format for evaluation purposes was seen as acceptable by both students and most music teachers, and the close correlation observed in the study between SOLO levels and Schools Board levels reinforced the validity of its use for this purpose.

The main intention of the test was to evaluate how well students understood what they heard. The test was not designed to measure the student's ability to hear, and despite the use of the Elements Checklist there was no intention, as in previous studies of student responses to music, to judge responses by means of the type of aspects of the music mentioned. The questions were deliberately designed to be as open as possible, so that the selection of musical elements was left entirely to the student. The expectation was that they would select those that they felt were most important in each piece, and despite there being no requirement to display an affective reaction to the music and the pretest mini-lesson having in fact stressed a cognitive objective approach, many students still chose to use associative and emotional ideas in their responses.

Listening in Schools Board Courses

It was initially envisaged that the procedures used in this

study might perhaps be considered as the basis of a moderating instrument in the Listening area of the Schools Board of Tasmania Grade 9 and 10 Music course. The introduction of the Tasmanian Certificate of Education courses with compulsory assessment criteria instead of broad course areas has made this idea irrelevant. Despite the objectives for these new courses including understanding and making informed judgments about music (cited in Chapter One) which are Relational level objectived, the sole Listening assessment criteria *"identifies and describes the main characteristics* of music" seems to be aimed at the Multistructural level. In this study the majority of responses were observed to be at this level, thus this emphasis does seem appropriate. However as 62% of responses from students achieving at the highest School Board level on the previous syllabus were classified at the Relational level, it may be that this criteria underestimates the ability of students of this age.

Applications in the Music Classroom

This study has shown that it is possible to use the SOLO Taxonomy as a general classification tool for the evaluation of Music Listening responses in the classroom. The SOLO Taxonomy classifies student responses into broad categories, it does not provide fine gradations between students, nor as was shown for the top students in the Retest Experiment, does it have the facility, in its present form, to indicate slight improvements. By itself it does not give sufficient information upon which to base a comparison between individual students or establish a detailed ranking system, so its main use with the present format of its evaluation system would be as a formative tool. As was noted in the Pilot Study and in work by Collis, Romberg and Jurdak (1986) tasks requiring skills at too high a level have a negative effect upon motivation towards further tasks, so the SOLO tests facility for sorting students into broad groups could be used to maintain motivation by establishing the level at which tasks should be set for each individual.

Within this study wide variations in the standard of responses at the Multistructural and Relational levels were noted. This was

acknowledged in the attempt to provide a scoring system mentioned in Chapter Four and in the feedback given to the Grade 9 and 10 students which was explained in Chapter Five. Although a classification system with categories within these levels was devised, it was not as reliable in use as the overall SOLO classifications, and so was not used in the analysis of results. It did however show that finer gradations were possible and could be made if needed in a summative evaluation situation. As it was felt that in classroom usage a teacher would use SOLO for broad classifications and then make finer distinctions based upon the types of musical elements expected from the current teaching emphasis or the particular piece of music, the construction of a more elaborate classification system for Music Listening was considered unnecessary, however the Elements Checklist could be used as the basis for one.

Every piece of music features different musical elements and uses them in unique ways, so although the Elements Checklist was useful as a preliminary guide and indicator of the quantity of ideas thought about; the mentioning of many elements would not be a measure of quality of listening if the elements mentioned were not those that were more important in this piece than in others. Before the checklist can be used effectively, therefore, the teacher must rate the importance of each element for each piece as a response to the specific task set. The checklist can then be used in conjunction with the overall SOLO levels to provide a detailed assessment of the students attainment on that task for that piece of music.

It has been shown how the SOLO Taxonomy can be adapted to provide an evaluation procedure for responses to music, and there have also been some indications that it might be possible to extend the technique for other classroom purposes. If the tasks set are carefully structured the technique could also probably be used to identify musical elements that were not properly understood or recognised by students, and it would therefore become a diagnostic tool for Music Listening.

Application to Music Teaching

Explaining the simple principles behind the SOLO levels could probably also be used as part of a teaching scheme similar to that used in the Retest Experiment which set out to improve student performances on Listening tasks. From discussion with the students in this Retest class, it seemed that the tuition given had not increased their ability to hear, but had made them more aware of the importance and relationships between various aspects of the music. This would seem to support the view that since Music is continually around us in our society, we all have the ability to hear it, but have to be taught to recognise and appreciate it. It could also be an indication that teaching students to use the SOLO levels could also be beneficial.

Teachers could also use the SOLO levels to structure the way they present Listening tasks. Bloom's Taxonomy was designed to do this, but due to the complexity of its categories and subcategories, it has not been extensively used in the classroom except in text books. The SOLO Taxonomy is a much simpler classification system, and it should therefore be easier to implement in the teaching situation. Questions such as *"What?"* and *"When?"* may be easily answered by students stating facts and their quick success may motivate students, but these questions do not stimulate students towards the higher levels of judgmental and speculative thinking as well as harder questions like *"Why?, Why not?, How?"* and *"What if?"*

There is a need for both types of question in the classroom, and perhaps part of the teacher's role is to provide opportunities for both and to maintain a balance between them; to demand statements of fact, and stimulate thought as to why those facts have been juxtaposed into a specific relationship; to show the relationship between active listening and questioning; to assist students to recognise the links between what is heard and the emotional impact it has on the listener. Music, an aural phenomena with a strong emotional impact, seems ideally suited to the teaching of relationships between ideas, and to developing an understanding of the relationship between the cognitive and affective domains.

It is also possible, though not shown in this study, that SOLO could also have a role to play in the development of original creative thinking in the classroom. The Taxonomy provides a framework into which students can place their separate ideas, and if tasks are clearly structured, use of the SOLO Taxonomy could lead them to recognise relationships and speculate about the possible effect of alterations which could be caused by minor changes. There seems to be ample scope for more research projects examining the application of the SOLO Taxonomy to Music teaching, which could be beneficial for all aspects of Music classroom practice.

Further Research Areas

Perhaps an extension to the present study should be undertaken. As the sample used in this study was limited to secondary school students in Grades 7-10, the results cannot be legitimately applied to students outside this age range or population base before further studies have been undertaken. The majority of responses obtained from this study were at the Multistructural level and it may be that this pattern would not be true for younger or older students. It might be necessary to alter the test format to an interview with younger students, which would make the technique less useful as a classroom assessment tool, but would be closer to a natural situation for younger students.

An extension of this study designed to investigate the higher SOLO levels would also be valuable for teachers at the college and tertiary levels. There were many Relational responses produced by the older students but there were no Extended Abstract responses given either under the test conditions or in the observed lessons during the Pilot Study. There was however some occasional evidence of thought at this level through references to material outside the stimulus and suppositions about the intentions behind the composer's use of some musical elements. It could be that the test time limit was too short, or that the questions set did not encourage speculation enough to stimulate responses at this level, or it might be that either older students or those with more musical training or

more motivation would produce responses completely at this level.

It would also be useful to extend the study with younger and older students, to ascertain whether the majority of responses were still Multistructural, or whether younger students would follow the trends noted by Collis, Romberg and Jurdak (1986) in their mathematical problem solving study. Investigations to see if there were differences in other age groups between those with and without specific musical training, to investigate whether Extended Abstract responses could be obtained, and to collect responses to other types of music and questions, could also provide valuable information for classroom music teachers.

The questions used in the Main Study were designed to be as open as possible, within the constraints imposed by the choice of music. Other types of questions, such as *" What does this music mean to you?"* which was used by Yingling in his 1962 study of college students, whilst it would obviously stimulate more affective responses should also encourage students to mention similar musical experiences, and might possibly lead to thought at the Extended Abstract level. Although in this study the question set for the second piece, which indicated clearly that a response which linked ideas together at the Relational level was expected, did not stimulate significantly more Relational responses; it could be that, when combined with tuition, this type of question could promote more Relational responses with this age group than other types of questions.

The possibility that tuition could improve SOLO levels which was indicated in the brief Retest Experiment, could also be investigated further. The students in this study reported that after the tuition, which was minimal, they felt better able to discuss pieces of music, and their Posttest levels were in many cases much higher than their Pretest levels, demonstrating that their feelings were correct. As the experiment was carried out at the end of the school year, it was not possible to see if this was only a short term effect or if they really had improved their ability to discuss what they had heard. It would be not only interesting but beneficial

to music teachers to repeat this experiment with another group of students at the beginning of their Grade 9 course and monitor their responses to music encountered throughout the year, and if possible extend this into a longtitudinal study.

It should also be possible to replicate studies such as those carried out by Gardner (1973), Castell (1982), and Hargreaves (1982) on style matching and analyse the responses obtained in terms of the SOLO Taxonomy instead of solely by the types of musical elements which they contain. The responses to this type of task which were observed by Gardner and quoted in the report on his study, could, as was noted in Chapter 3, be classified into SOLO levels. It might be however that the entire picture might be more complex than these few brief quotations indicated, and that further research would show this.

Application to Other Areas of Music Education

If SOLO can be applied to the evaluation of Music Listening as has been shown, then it seems reasonable to surmise that it could perhaps also be applied to the other areas of the music curriculum: Performance and Composition.

A common Performance problem, the student who thinks that they have mastered a piece of music because they can play the right notes, but who is not playing the correct rhythm with these notes, and has not begun to add slight alterations in volume, phrasing, articulation, and timbre, which would turn the notes into a piece of music, could for instance be analysed in terms of the SOLO Taxonomy. The Unistructural approach would be to focus on one aspect of the music, the pitch or rhythm or harmony, whereas combining these elements could indicate the Multistructural level. Using other elements to realise the musical or emotional nature of the piece could be thought of as being a Relational level skill, whereas creating a fresh interpretation would be a demonstration of Extended Abstract level functioning. The restricted Unistructural and Multistructural attitudes towards musical performance can often be heard in the synthesised music which is provided by telephone

companies to be played whilst incoming calls are placed on hold. In these pieces the synthesizer has been programmed to play the correct notes at the correct time, but the result is not recognisable or acceptable as music due to the lack of phrasing, metrical stresses, variations in timbre, dynamics or articulation which are expected in aesthetically satisfying performances.

There is also a possibility that musical compositions could also be analysed and evaluated in a similiar way, with broad classifications being made on the basis of SOLO levels, followed by a detailed investigation of the efficacy of the use of each element of the music. It might also be possible to observe the stages students pass through whilst learning or composing a piece of music and analyse these using SOLO. It would be interesting to compare a SOLO based study of Composition with that carried out by Swanwick and Tillman (1986), and the results of studies such as these would be extremely useful to music teachers both for assessment and teaching purposes.

It should also be possible to analyse tasks set in the music classroom in terms of the SOLO Taxonomy, so that the relative complexities of different tasks could be established. Music reading, for instance, a major hurdle for many students, noticeably in Grade 8, is a task requiring the coordination of knowledge about pitch and rhythm notation, musical signs and terms, and instrumental fingering with the ability to produce and control the sounds on an instrument. It is undoubtedly a complex Relational skill, and this could account for the problems in this area encountered by many students. If teachers are aware of the essentially Relational nature of tasks like this, then they can break their teaching down into Unistructural steps, combine these into Multistructural exercises and lead the more able students into tackling the more complex Relational tasks by themselves, whilst providing simpler routes to the same goals for the other students. Awareness of complexities like this in the music curriculum would be of great assistance to teachers both in planning their programmes and daily lessons, and in dealing with students having difficulties with tasks set.

If the pattern of levels observed in this study of Music Listening, where the majority of responses in all grades were at the Multistructural level, holds true for the other areas of the music curriculum, then it could be that Grade 8 is too early for the introduction of many Relational level tasks. It might also be that the decline in motivation towards Music in Grade 8, which was mentioned by many of the teachers contributing to this study, could be partially created by the introduction of too many tasks requiring Relational thinking at this age.

Further studies examining motivation towards Music lessons at different ages and between sexes, determining which SOLO levels can be expected for musical tasks in different grades, and analysing the tasks required in music curricula, could not only provide much useful information for music teachers, but also contribute to the development of effective teaching strategies. These are all problems which have not been approached in the present study, but which have been stimulated by it.

Summary

The present study, which began as an attempt to devise an effective evaluation tool for use in moderating between students in the same school and possibly between schools, has developed an evaluation technique that can be used with an entire class, uses actual pieces of music rather than artificial sounds, is simple to administer, and is acceptable to both students and teachers. Reliability of the test has been shown through inter- and intra-judge assessments of response levels, and through the stability of response levels across different pieces and tasks and, in the absence of changes in music tuition, across retests of the same students. The SOLO test has also been shown to be measuring what was required by the Schools Board syllabus without undue influences being made upon response levels by factors other than the amount of music tuition being received at the time of the test, which demonstrates the validity of its use for assessment purposes.

The test has also been shown to have possibilities for

extension to other pieces of music, types of questions and listening situations such as solo, ensemble, composition or arrangement rehearsals or live performances, and there seem to be some possibilities for its application to other music evaluation situations. As the basic principles behind the SOLO Taxonomy have also been shown to be easily grasped by students and teachers it may also have some applications in the areas of self-evaluation and teaching.

Music Education is a process of developing skills and understanding in the three areas of Performance, Listening and Composition. However despite being a process based subject, there is still a requirement that the end products of each area be evaluated. Evaluation of the products created during the learning process will reveal the efficiency of the teaching process, identify aspects which need reteaching and relearning, and assist planning for future lessons, so there are needs for quick and effective devices for regular use for formative evaluation as well as for summative evaluation. The evaluation technique developed in this study can be used to evaluate the quality of progress at any moment in the process, and so can assist in the improvement of music education.

Bibliography

References

ANDREWS,M.L.& MADEIRA, S.S. (1977).
The Assessment of Pitch Discrimination Ability in Young Children.
Journal of Speech and Hearing Disorders, 42, 279-286.

ASMUS,E.P.(1985).
The Development of a Multidimensional Instrument for the Measurement of Affective Responses to Music.
Psychology of Music, 13(1), 19-30.

BAMBERGER,J.S., & BROFSKY,H. (1979).
The Art of Listening : Developing Musical Perception.
New York : Harper and Row

BARTLETT, D.L. (1973).
Effect of Repeated Listenings on Structural Discrimination and Affective Response.
Journal of Research in Music Education, 21 (4), 302-317.

BENTLEY,A.(1966).
Measures of Musical Ability.
New York : October House.

BIGGS,J.B.& COLLIS,K.F (1982).
Evaluating the Quality of Learning : The SOLO Taxonomy (Structure of the Observed Learning Outcome).
New York : Academic Press.

BLOOM,B.S.,ENGELHART,M.D.,FURST,E.J.,HILL,W.H.,& KRATHWOHL,D.R.(1956).
Taxonomy of Educational Objectives : Cognitive Domain.
New York : McKay

BLOOM,B.S.,HASTINGS,J.T.,MADAUS,G.F. (1971).
Handbook of Formative and Summative Evaluation of Student Learning.
New York : McGraw Hill

BOYLE,J.D. & RADOCY,R.E. (1982).
Music Education. In Mitzel,H. (Ed) *Encyclopedia of Educational Research. 5th Edition.* New York : Free Press (Macmillan)

BRADLEY,I.L (1971).
Repetition as a Factor in the Development of Musical Preferences.
Journal of Research in Music Education, 19 (3), 294-298.

BRADLEY,I. (1972).
Effect on Student Musical Preferences of a Listening Program in Contemporary Music.
Journal of Research in Music Education, 20, 344-353
BRIDGES,D. (1974).
Australian Test for Advanced Music Studies. Melbourne, ACER.
BRIDGES,V. (1965).
An Exploratory Study of the Harmonic Discrimination Ability of Children in Kindergarten through Grade Three in Two Selected Schools. Unpublished Doctoral Dissertation, Ohio State University.
BULA,K., & SZYMANOWSKI,K. (1987).
The Participation of the Verbal Factor in Perception of Musical Compositions.
Bulletin of the Council for Research in Music Education, 91, 15-18
CAAMA MUSIC RECORDS AND CASSETTES (1987).
Bushfire Radio.
Alice Springs : Central Australian Aboriginal Media Association.
CAAMA (CENTRAL AUSTRALIAN ABORIGINAL MEDIA ASSOCIATION) (1989).
The CAAMA Group : An Introduction to the Caama Group of Companies.
Alice Springs : Central Australian Aboriginal Media Association.
CASE,R. (1980).
The Underlying Mechanism of Intellectual Development. In Kirby & Biggs (Eds) *Cognition, Development and Instruction.*
New York : Academic Press
CASE,R. (1985).
Intellectual Development : Birth to Adulthood.
Orlando, Fla. : Academic Press
CASTELL,K.C. (1982).
Childrens' Sensitivity to Stylistic Differences in Classical and Popular Music.
Psychology of Music, Special Edition, 22-25.
COLLES,H.C. (1919).
The Growth of Music.
Oxford : Clarendon Press
COLLIS,K.F, ROMBERG,T.A, JURDAK,M.E (1986)
A Technique for Assessing Mathematical Problem-Solving Ability
Journal for Research in Mathematics Education, 17,3 (206-221).

COLWELL,R.(1968-69).
Music Achievement Test.
Cited in Colwell,R. (1970).
COLWELL,R. (1970)
The Evaluation of Music Teaching and Learning.
New Jersey : Prentice Hall, 148-150.
COLWELL,R. (1987).
Test Item Difficulty and Perception.
Bulletin of the Council for Research in Music Education, 91, 19-22.
CONLEY,J.K (1981).
Physical Correlates of the Judged Complexity of Music by
Students Differing in Musical Background.
British Journal of Psychology, 72, 451-464.
COPLAND,A. (1957).
What To Listen For In Music.
New York : McGraw Hill.
CRICKMORE,L. (1968).
An Approach to the Measurement of Music Appreciation.
Journal of Research in Music Education, 16 (4), 239-253.
CRUMMER,G.C.,HANTZ,E.,CHUANG,S.W.,WALTON,J. & FRISINA,R.D. (1988)
The Neural Basis for Music Cognition : Initial Experimental Findings.
Psychomusicology, 7 (2), 117-125.
DONALDSON,M. (1978).
Children's Minds.
Glasgow : Fontana/Collins
DOROW,L.G. (1977).
The Effect of Teachers Approval/Disapproval Ratios on
Student Music Selection and Concert Attentiveness.
Journal of Research in Music Education, 25 (1), 32-40.
EDWORTHY,J. (1985).
Melodic Contour and Musical Structure.
In Howell, Cross & West (Ed) (1985)
FARNSWORTH (1954). cited in Farnsworth (1969)
FARNSWORTH,P.R. (1969).
The Social Psychology of Music, 2nd Edition.
Ames, Iowa : Iowa State University Press

FARNUM,S.E. (1969).
Farnum String Scale.
Cited in COLWELL,R. (1970).
FISCHER,K.W. (1980).
A Theory of Cognitive Development :
The Control and Construction of Hierarchies of Skills.
Psychological Review, 87 (6), 477-531
FLOWERS,P.J. (1983).
The Effect of Instructions in Vocabulary and Listening on
Nonmusicians Descriptions of Changes in Music.
Journal of Research in Music Education, 31 (3), 179-189.
GABRIEL,C. (1978).
An Experimental Study of Deryck Cooke's Theory of Music and Meaning.
Psychology of Music, 6 (1), 13-20
GALWAY, J. (1986).
My Life as a Flute Player. ABC Radio Interview.
GARDNER,H.(1973).
Children's Sensitivity to Musical Styles.
Merrill Palmer Quarterly, 19, 67-77.
GARDNER,H.(1981).
Do Babies Sing A Universal Song ?
Psychology Today, 15 (12), 70-76
GARDNER,H. (1984).
Frames of Mind : The Theory of Multiple Intelligences.
London : Heinemann
GERINGER,J.M. (1982).
Verbal and Operant Music Listening Preferences in
Relationship to Ages and Musical Training.
Psychology of Music, Special Edition, 47-50.
GETZ,R.P. (1966).
The Effects of Repetition on Listening Response.
Journal of Research in Music Education, 4, 178-192.
GINSBURG,H., & OPPER,S. (1969).
Piaget's Theory of Intellectual Development.
New York : Prentice Hall

GOOLSBY,T.W (1984).
Concepts and Skills for the Appreciation of Music.
Journal of Aesthetic Education 18 (4), 15-33
GORDON,E. (1965).
Musical Aptitude Profile.
Boston, Mass. : Houghton Mifflin.
GREER,D.,DOROW,L.G.,WACHHAUS,G.,& WHITE,E. (1973).
Adult Approval and Students' Music Selection Behavior.
Journal of Research in Music Education, 21 (4), 345-354.
GREER,D.,DOUGLAS.R.,DOROW,L., & HANSER,S. (1973).
Music Discrimination Training and the Music Selection Behaviour of Nursery and Primary Level Children.
Bulletin of the Council for Research in Music Education, 35, 30-43.
HAIR,H. (1977).
Discrimination of Tonal Direction on Verbal and Nonverbal Tasks by First Grade Children.
Journal of Research in Music Education, 25 (3), 197-210.
HAIR,H. (1981).
Verbal Identification of Musical Concepts.
Journal of Research in Music Education, 29 (1),11-21.
HAIR,H.I. (1987).
Descriptive Vocabulary and Visual Choices : Children's Responses to Conceptual Changes in Music.
Bulletin of the Council for Research in Music Education, 91, 59-64
HARGREAVES,D.J. (1982).
The Development of Aesthetic Reactions to Music.
Psychology of Music, Special Edition, 50-54.
HARGREAVES,D.J. & CASTELL,K.C. (1987).
Development of Liking for Familiar and Unfamiliar Melodies.
Bulletin of the Council for Research in Music Education, 91, 65-69.
HARGREAVES,D.J. & COLMAN,A.M. (1981).
The Dimensions of Aesthetic Reactions to Music.
Psychology of Music, 9 (1), 15-20.
HARROW,A.J. (1972).
A Taxonomy of the Psychomotor Domain : a Guide for Developing Behavioral Objectives.
New York : McKay

HARVEY,A. (1986).
Is Brain Research Relevant for Music Education?
British Journal of Music Education, 3 (2), 175-179.
HEDDEN,S.K. (1981).
Music Listening Skills and Music Listening Preferences.
Bulletin of the Council for Research in Music Education, 65, 16-26.
HERBERGER,R. (1983).
Presenting a Method of Analysing Mental and Emotional Processes in Secondary School Students While They Are Listening To Music.
Bulletin of the Council for Research in Music Education, 75, 40-47.
HERBERGER,R. (1987).
The Degree of Attractiveness to 15 Year Old School Students in the GDR of Different Styles, Genres and Trends of Contemporary Music - Results of a Factor Analysis.
Bulletin of the Council for Research in Music Education, 91, 70-76.
HEVNER,K. (1935).
Expression in Music : A Discussion of Experimental Studies and Theories.
Psychological Review, 42, 186-204.
HEVNER,K.M. (1956).
Studies in Music Appreciation.
Journal of Research in Music Education, 4, 3-25.
HEYDUK,R.G. (1975).
Rated Preference for Musical Compositions as it Relates to Complexity and Exposure Frequency.
Perception and Psychophysics, 17 (1), 84-91.
HILL,J.D. (1968).
A Study of the Musical Achievement of Culturally Deprived Children and Culturally Advantaged Children at the Elementary School Level.
Unpublished Doctoral Dissertation, University of Kansas.
HOPKINS,A. (1979.)
Understanding Music.
London : Dent.
HOWELL,P.,CROSS,I., & WEST,R. (Eds) (1985).
Musical Structure and Cognition.
Academic Press, London

HUFSTADTER,R.A. (1977).
An Investigation of a Learning Sequence of Music Listening Skills.
Journal of Research in Music Education, 25 (3),184-196.
HULSE,S.H. (1990)
The Acquisition of Pitch and Rhythm in Songbirds.
In Wilson,F.R & Roehmann F.L. (Eds.)
Music and Child Development : Proceedings of the 1987 Denver Conference, The Biology *of Music Making.*
Missouri : MMB Music
JETTER,J.T. (1978).
An Instructional Model for Teaching Identification and Naming of Music Phenomena to Preschool Children.
Journal of Research in Music Education, 26 (2), 21-31.
JORGENSEN,E.R. (1987).
Percy Scholes on Music Appreciation : Another View
British Journal of Music Education, 4 (2), 139-159
KRATHWOHL.D.R.,BLOOM,B.S.,& MASIA,B.B. (1964)
Taxonomy of Educational Objectives : The Affective Domain.
New York : McKay
KUHN,T.L. (1974).
Discrimination of Modulated Beat Tempo by Professional Musicians.
Journal of Research in Music Education, 22 (4), 270-277.
KWALWASSER J.,RUCH G.M.(1927, 1952).
Kwalwasser-Ruch Test of Musical Accomplishment for Grades Four Through Twelve.
Cited in COLWELL,R.(1970).
LAMBORN,S.D., & FISCHER,K.W. (1988).
Optimal and Functional Levels in Cognitive Development : The Individual's Developmental Range.
Newsletter : Int. Soc. for the Study of Behavioural Development No 2, Serial No 14, 1-4
LARSEN,R.L. & BOODY,C.G. (1971).
Some Implications for Music Education in the Work of Jean Piaget.
Journal of Research in Music Education, 19 (1), 35-50.
LEASK.J. (1989).
Upbeat : Music Education in the Classroom. Level Five Teacher's Book.
Gosford : Ashton Scholastic

MADSEN,C.K. (1979).
Modulated Beat Discrimination Among Musicians and Non Musicians.
Journal of Research in Music Education, 27 (2), 57-67.
MADSEN,C.K.& PRICKETT,C.A. (Eds) (1987).
Applications of Research in Music Behavior.
University of Alabama Press, Tuscaloosa
MCDONALD,D. (1974).
Environment : A Factor in Conceptual Listening Skills of
Elementary School Children.
Journal of Research in Music Education, 22 (3), 205-214.
MENUHIN,Y.& DAVIS,C.W (1979).
The Music of Man.
Toronto : Methuen.
MYERS,C.S. (1927).
Individual Differences in Listening To Music.
In Schoen,M. *The Effects of Music*, New York : Harcourt Brace.
NELSON,D.J. (1984).
Trends in the Aesthetic Responses of Children to the
Musical Experience.
Journal of Research in Music Education, 33 (3), 193-203.
NIERMAN,G.E. (1983).
The Effects of Grade Level on Secondary Music Students'
Perceptive-Descriptive Skills.
Psychology of Music, 11 (2), 73-78.
O'HEARN,R.N. (1984)
An Investigation of the Response to Change in Music Events
by Children in Grades One, Three and Five.
Unpublished Ph.D thesis.
OLSON,I. (1984).
Measurement of Musical Awareness.
Bulletin of the Council for Research in Music Education, 77, 31-42.
ORTMANN.O. (1927).
Types of Listeners : Genetic Considerations.
In Schoen,M. *The Effects of Music*, New York : Harcourt Brace.
PARSONS,M.,JOHNSTON,M., & DURHAM,R. (1978).
Developmental Stages in Children's Aesthetic Responses.
Journal of Aesthetic Education, 12 (1), 83-104

PARSONS,M.J. (1987).
How We Understand Art :
A Cognitive Developmental Account of Aesthetic Experience.
New York : Cambridge University Press
PAYNE,E. (1980).
Towards an Understanding of Music Appreciation.
Psychology of Music, 8 (2), 30-41.
PAYNTER,J. (1982).
Music in the Secondary School Curriculum.
Cambridge : Cambridge University Press.
PICK,A.,PALMER,C.,HENNESSY,B.,UNZE,M.,JONES,R. & RICHARDSON,R.M.(1988).
Children's Perception of Certain Musical Properties :
Scale and Contour.
Journal of Experimental Psychology 45, 28-51.
PRINCE,W.F. (1972).
A Paradigm for Research on Music Listening.
Journal of Research in Music Education, 20 (4), 445-454.
RADOCY,R.E. (1982).
Preference for Classical Music : A Test of the Hedgehog.
Psychology of Music, Special Edition, 91-95.
ROSENSTIEGEL,A.K.,MORISON,P.,SILVERMAN,J., & GARDNER,H. (1978).
Critical Judgment : A Developmental Study.
Journal of Aesthetic Education, 12 (4), 95-107
SADIE,S., & LATHAM,A. (Ed) (1985).
The Cambridge Music Guide.
Cambridge : Cambridge University Press
SEASHORE, C.E. (1938).
Psychology of Music.
New York : McGraw Hill.
SERAFINE,M.L. (1980).
Piagetian Research in Music.
Bulletin of the Council for Research in Music Education, 1-21.
SHETLER,D.J. (1990)
The Inquiry into Prenatal Musical Experience
In Wilson,F.R & Roehmann F.L. (Eds.) *(1990)*

SIMON,C.R.& WOHLWILL,J.F. (1968).
An Experimental Study of the Role of Expectation and Variation in Music.
Journal of Research in Music Education, 16 (3), 227-237.
SIMONS,G. (1976).
A Criterion Referenced Test of Fundamental Music Listening Skills.
Child Study Journal, 6 (4)
STANDIFER,J.A. (1970).
Effects on Aesthetic Sensitivity of Developing Perception of Musical Expressiveness.
Journal of Research in Music Education, 18, 112-125.
STECK,L.& MACHOTKA,P. (1975).
Preference for Musical Complexity : Effects of Context.
Journal of Experimental Psychology : Human Perception and Performance. 104 (2), 170-174.
STOWASSER,H. (1989).
Discover Music Making.
Melbourne : Longman Cheshire
STURGES & MARTIN (1974).
Rhythmic Structure in Auditory Temporal Pattern Perception and Immediate Memory.
Journal of Experimental Psychology, 102 (3), 377-383
SWANWICK,K. (1973)
Music Cognition and Aesthetic Response.
Psychology of Music. 1 (2), 7-13.
SWANWICK,K. (1975).
Can There Be Objectivity in Listening to Music ?
Psychology of Music. 3 (2), 17-23.*33*
SWANWICK,K. (1982).
Discovering Music : Developing the Music Curriculum in Secondary Schools.
London : Batsford Academic and Educational.
SWANWICK,K., & TILLMAN,J. (1986).
The Sequence of Musical Development :
A Study of Children's Compositions.
British Journal of Music Education, 3 (3), 305-339

SWANWICK,K. (1988a).
Music, Mind and Education.
London : Routledge.
SWANWICK,K. (1988b).
The Relevance of Research : Too Little Theory.
In Kemp,A.(Ed) (1988)
SWICKARD, J.H. (1971).
A Comparative Study of Musical Achievement of Students in Grades Four, Five and Six.
Unpublished Doctoral Dissertation, University of Illinois.
TILLMAN,J. (1989)
Towards a Model of Childrens' Musical Creativity.
Canadian Music Educator. 30 Research Edition, 169-174.
THOMSON,W. (1978).
Music for Listeners.
Englewood Cliffs, New Jersey : Prentice Hall
VALOVY,E. (1981).
Problems of Research of an Objective Character into the Meanings and Context of a Musical Work.
Bulletin of the Council for Research in Music Education, 66/67, 92-95.
VYGOTSKY,L.S. (1978).
Mind in Society : the Development of Higher Psychological Processes.
Cambridge,Mass. : Harvard University Press
WATKINS FARNUM (1954). Performance Scale.
Cited in Colwell,R. (1970).
WAGNER,M.J.& HARDING,B.J. (1987).
The Effect of Guided Listening Information and Music on the Alpha Brainwave Production of Musicians and Non Musicians.
In Madsen & Prickett,C.A. (Eds) (1987)
WEST,R.,HOWELL,P.,CROSS,I. (1985).
Modelling Perceived Musical Structure.
In Howell,P.,Cross,I., & West,R. (Eds) (1985).
WILSON,F.R. (1985).
Music As Basic Schooling For The Brain.
Music Educator's Journal, May 1985 39-42

WILSON, F.R (1988).
Music and Your Child.
St. Louis : A.M.C., 1-2
WOODS,D.G.(1990).
Musical Creativity and Brain Functioning.
Paper presented at the Australian Society for Music Education Conference, Alice Springs, July 1990.
WOODERSON,D.C. & SMALL,A.R. (1981).
Instrument Association Skills : Children in 1st and 2nd Grades.
Journal of Research in Music Education, 29 (1), 39-46.
YINGLING,R.W. (1962).
Classification of Response Patterns in Listening To Music.
Journal of Research in Music Education, 10, 105-120.
ZENATTI,A. (1976).
Melodic Memory Tests : A Comparison of Normal Children and Mental Defectives.
Journal of Research in Music Education, 23, 41-52.

Reading List

ABELSON,P.H.,BUTZ,E.,& SNYDER,S.H.(Eds) (1985)
Neuroscience.
Washington,D.C. : American Association for the Advancement of Science.
BALL,W.A. (1988)
Expressing the Inexpressible : Developing An Aesthetic Vocabulary.
Music Educators Journal, April, 53-56.
BARTHOLOMEW,D. (1987).
Problems with Piagetian Conservation and Musical Objects.
Bulletin of the Council for Research in Music Education, 93, 25-40.
BENTLEY,A. (1966).
Musical Ability in Children and Its Measurement.
London : Harrap.
CAMPBELL, D.G. (1983).
Introduction to the Human Brain.
Saint Louis, Mo. : MMB Music.

CARLSEN,J.C.,LADEN,B. (1988).
Mental Representations in Music
In Kemp,A. (Ed) (1988)
CLYNES, M. (Ed) (1982).
Music, Mind and Brain : the Neuropsychology of Music.
New York : Plenum Press.
CRITCHLEY,M.,HENSON,R.A. (Eds). (1977).
Music and the Brain.
London : Heinemann.
DELIS,D.,FLEER,J.,KERR,N.H. (1978).
Memory for Music.
Perception and Psychophysics 23, 251-218.
DEUTSCH,D. (1982).
The Psychology of Music.
New York : Academic Press.
DEUTSCH,D.,DEUTSCH,J.A. (1975).
Short Term Memory.
New York : Academic Press.
HAACK,P. (1969).
A Study in the Development of Music Listening Skills of
Secondary School Students.
Journal of Research in Music Education 17, 193-201.
HEDDEN,S.K (1980).
Development of Listening Skills.
Bulletin of the Council for Research in Music Education, 12-22.
HELLER,J.,CAMPBELL,W. (1981).
A Theoretical Model of Music Perception and Talent.
Bulletin of the Council for Research in Music Education. 66/67, 20-24
HELLER,J.,CAMPBELL,W. (1982)
Music Communication and Cognition.
Bulletin of the Council for Research in Music Education, 72, 1-15
HELLER,J.,CAMPBELL,W.,GIBSON,B. (1982).
The Development of Music Listening Skills in Children.
Psychology of Music, Special Edition, 55-58.
KEMP,A. (Ed) (1988).
Research in Music Education : A Festschrift for Arnold Bentley.
Trowbridge, Wiltshire : International Society for Music Education

LANGER,S.K. (1949)
Philosophy in a New Key.
Cambridge, Mass. : Harvard University Press
MADSEN,C.K & YARBOROUGH,C. (1980)
Competency Based Music Education.
Englewood Cliffs, New Jersey : Prentice Hall
MEYER,L.B. (1961).
Emotion and Meaning In Music.
Chicago : University of Chicago Press.
PRINCE,W. (1977).
Measuring Discrimination of Complex Musical Events.
Journal of Research in Music Education, 25 (2), 91-99.
REIMER,B. (1970).
A Philosophy of Music Education.
Englewood Cliffs : Prentice Hall.
SCHOEN,M. (1944).
The Enjoyment of the Arts.
New York : Philosophical Library.
SERAFINE, M.L. (1988).
Music As Cognition : The Development of Thought in Sound.
New York : Columbia University Press.
SHUTER,R. (1968).
Psychology of Musical Ability.
London : Methuen.
SLOBODA, J.A. (1986).
The Musical Mind : The Cognitive Psychology of Music.
Oxford : Clarendon Press.
WILSON,F.R. & ROEHMANN,F.L.(Eds) (1990)
Music and Child Development :
Proceedings of the 1987 Denver Conference, The Biology of Music Making.
St Louis : MMB Music
ZIMMERMAN,M.P. (1984).
The Relevance of Piagetian Theory for Music Education.
International Journal of Music Education. 31-34.

Appendix A

Tasmanian Education and the Schools Board of Tasmania

Outline of the Tasmanian Education System

In Tasmania compulsory education is carried out in Kindergartens for students aged 5, Primary Schools from ages 6-11 and in 4 year High Schools from ages 12-16. Students may leave school aged 16 or continue their academic or vocational education in Colleges. The Year groups are known as Grades. Grades 1- 6 are found in Primary Schools, Grades 7-10 in High Schools and Grades 11-12 in Colleges.

Music is taught in all schools and colleges, usually in heterogeneous classes from Grades 1-8, and in elective general Music classes in Grades 9-12.

In Primary Schools Music may be taught by the classroom teacher or by itinerant specialists who are often not available in rural areas. Instrumental work is usually taught by the same teacher, a parent, or in a few urban schools by itinerant instrumental teachers as part of special Primary Band or String Teaching programmes. The official Primary Music Guidelines focus on Music Making and Creating using voices and instruments.

In High Schools a general Music course is taught in Grades 7 and 8 by a specialist Music teacher with, in urban areas, support from itinerant instrumental teachers. The Music course includes music reading, listening, general knowledge, composition, improvisation and instrumental tuition leading to the formation of concert, brass and rock bands, and in some schools choirs and string ensembles. In Grades 9 and 10 Music becomes an elective subject certificated by the Schools Board of Tasmania, and known as a School Certificate course.

The School Certificate Grade 9 and 10 Music Course at the time of this study was a two year course, with equal value given to Performance both solo and ensemble, Composition including improvisation and arranging skills, and Listening including aspects of ethnic and western musics. The TCE (Tasmanian Certificate of Education) which was introduced in Grade 9 in 1990, has introduced drastic changes. All courses are now one-year courses with certification beginning in Grade 9. In Grades 9 and 10 the new courses cover approximately the same content with assessment being through 11 or 12 compulsory criteria.

At grade 11 and 12, College level, Music has only been available in the form of one year practical or pre-tertiary academic courses, known as Higher School Certificate Courses. The one year format of these courses will remain under the TCE.

Schools Board of Tasmania Award System

This system was in operation until the end of 1990, when it was superceded by the TCE. The School Certificate had three levels of syllabus. Level One was the easiest level, designed for students with some learning problems. Rote learning of music for performance and simple tasks were features of this level. Level Two was the syllabus aimed at the majority of students, and Level Three was the most difficult syllabus, aimed at the more academic students. At the Higher School Certificate level, Level Three subjects were required for University entrance.

At the School Certificate level (Grade 10) there were 5 awards: Credit, Higher Pass, Pass, Lower Pass and No award (failure). Awards were usually written as 3C (Level Three Credit), 2H (Level Two Higher Pass) or 1L (Level One Lower Pass). Students were introduced to this system in Grade 8 or 9.

The TCE courses for Grades 9 and 10 also have three parallel syllabi of varying degrees of difficulty, which are designed to be taught in the same classroom and have many similarities in content.

Feedback using Schools Board Awards in the Pilot Study

During the year of the Pilot Study all the Grade 10 Listening tasks were assessed using SOLO, but the students were given feedback in the form of the Schools Board Awards. This was to give tham progress information in a form they understood without alerting them to the criteria behind the SOLO assessment. Figure A.1 shows how the two systems were linked.

Figure A.1
Schools Board Awards and SOLO Levels

Credit	*Fully Relational and Extended Abstract*
Higher Pass	*Multistructura/Relational Transitional*
Pass	*Multistructural*
Lower Pass	*Unistructural*
Fail (N)	*Prestructural*

Feedback for Grade 9 and 10 Students after testing sessions

In the Main Study feedback from the SOLO test was given using letter ratings as this was one of the methods used in the school for coursework assessment. The Schools Board awards were, in many subject areas, often only used for end of unit/course/term/year summative assessment. The explanation of the letter ratings which was published for students is on Figure A.2.

Figure A.2
Feedback Symbols and Explanations for Grades 9 and 10

A *An excellent answer. You mention many aspects of the music, explain your ideas clearly, and connect ideas and musical facts together efficiently.*
(Fully Relational)

3 *You heard a lot of the ideas the composer put into this piece of music, and are starting to explain musical ideas using the evidence you have found in the music.*
Well done, you are becoming an intelligent listener.
(Multistructural + Relational ideas,
Multistructural Transitional response)

C *You have heard a lot of the ideas the composer put into this piece of music. Good.*
(Multistructural)

D *You only mentioned one aspect of the music. Didn't you hear the others ? or didn't you know how to write them down ?*
I think you may need some help.
(Unistructural)

? *You didn't answer the question.*
Did you understand what to listen for ?
You may need some help next time.
(Prestructural)

Appendix B

Possible GCSE Grade-related Criteria for Listening

Grade G : SENSORY The student recognises clear differences of loudness level; widely different pitch differences; obvious changes of tone colour and texture. None of these is technically analysed and there is no account of expressive character or structural relationships.

Grade F : MANIPULATIVE The student identifies, but does not analyse, devices to do with the management of musical material; for example, trills; tremolos; scalic patterns; glissandi; steady or fluctuating beat; spatial and stereophonic effects; like and unlike instrumental sounds.

Grade E : PERSONAL EXPRESSIVENESSS The student describes the general atmosphere, mood or character of a passage and recognizes changes of expressive level, without drawing attention to structural relationships. Descriptions of the music may be in terms of dramatic incident, stories, personal associations and visual images, or feeling qualities.

Grade D : VERNACULAR The student recognizes common musical procedures and may identify such elements as metre, phrase shape and length, repetitions, syncopation, sequences, drones, ostinati. There is some technical analysis.

Grade C : SPECULATIVE The student identifies what is unusual or unexpected within the context of a particular work and is able to draw attention to changes of character by reference to instrumental or vocal colour, pitch, speech, loudness, rhythm and phrase length; the magnitude and frequency of changes, the extent to which changes are gradual or sudden.

Grade B : IDIOMATIC The student places music within a stylistic context and shows awareness of technical devices and the structural procedures that characterize a particular idiom; such as transformation by variation, decoration and contrasting middle sections, distinctive harmonies and rhythmic inflections, specific instrumental sound production or vocal melisma.

Grade A : SYMBOLIC In addition to meeting the criteria for the other grades, the student displays evidence of personal exploration and commitment through an account of a chosen area of musical investigation. There will be evidence of individual insights and sustained engagement with particular works, performers or composers.

Swanwick (1988a)

Appendix C

Assessment Procedures Examined In The Pilot Study

1 : Individual interviews

The use of oral questions interfered with the aural recall of the music. Interviews were also found to be too time consuming to be of use to the teacher of a class of over 25 students.

2 : Class Discussions

Teacher-led discussions are a normal part of the music teaching process, but they do not permit every student to express their thoughts. Students were influenced by what previous students had said and were reluctant to put forward ideas which were radically different for fear of ridicule. It was also almost impossible to monitor and record the views of each student, and ensure equal participation whilst maintaining class control especially during lively discussions, responses could however be classified according to SOLO.

3 : Group Discussions

In an attempt to create a situation where every student's voice could be heard, small groups of 4-6 students were set up. They discussed the music heard, shared their ideas and then reported back to the class. In carefully structured groups the unsure were provided with a supportive situation, though there was a tendency for domination by the more vocal. The reporting back stage was also sometimes dominated by a few and not an accurate reflection of the views of the entire group. It was impossible in this situation to monitor the contribution of individual students, so this method was not considered suitable for use as an assessment procedure, though it could be an excellent teaching strategy.

4 : Student Talks about Music They Have Listened To

Individuals listened to different pieces, which were chosen by themselves and then reported back to the class. This was a most revealing and rewarding classroom exercise, particularly when other students became interested in the pieces and asked pertinent questions. It did however consume vast amounts of class time, and if the reporting back sessions had not been spread over several lessons and interspersed with other activities, it could have become boring for the other students. It had several problems as an assessment technique used to compare students with each other. All students could not be given the same piece of music as a stimulus, so

comparisons would have been made on unequal material. Some pieces were selected that were well known and students were able to research information about them and find formal analyses and reviews written by professionals upon which they based their talks. In some cases this was well done, with students trying to understand the ideas they had read, examining these learned opinions in the light of their own listening, and providing personal commentaries and fresh thoughts on the pieces. These students were obviously working at a well integrated Relational level. Other students prepared themselves in similar ways but merely repeated what they had read without trying to understand it and used it to avoid the necessity of thinking for themselves. They failed to provide any evidence that they had in fact listened to the music. Giving a talk in front of the class was also a traumatic experience for some shy students, who felt they had not shown themselves to good advantage and requested permission to submit a written paper in addition. These written papers did give evidence of detailed listening which had often been an emotional and rewarding experience for the students. This procedure was therefore found to be suitable for use as an indicator of an individual student's ability to listen, observe and verbalise in public about what they hear, but not suitable for all students, and not suitable as a comparative measure between students.

Oral procedures having proved impractical for the purpose of assessment in a classroom setting, attention was then focussed on written procedures.

5 : Assignments

Assignments can be set which involve listening to music and commenting on aspects of the music. As they are usually set over a period of time, there is opportunity for collusion between students, discussion with other musicians and inclusion of material from printed and recorded sources; all of which are valuable additions to the learning process, but mean that the finished piece of work cannot be guaranteed to be the unaided work of the student.

6 : Worksheets

Worksheets can be administered quickly in class under conditions which ensure that all answers are the students' own work, or used in peer support situations and as the basis of groupwork. They can be constructed to require answers at various levels of knowledge and complexity, given comprehensive marking schemes and are recyclable. They have therefore become a valuable part of the normal teaching process, and their use seemed an obvious method for examining SOLO levels.

Superitem worksheets were therefore written in which the stimulus or stem was a piece of music which the students listened to together, and the questions required answers at the Unistructural, Multistructural and Relational levels. which could only be answered by listening to the music. An open-ended question was also included which could be answered satisfactorily at the Relational level, but which required a judgment to be made and insinuated that other musical knowledge was relevant and acceptable. It was hoped that

this might provide scope for those wanting to provide Extended Abstract answers, whilst not discouraging the other students.

These worksheets were presented to Year 7 and 8 students as part of their normal coursework. In the initial trial students worked on them alone, but many did not answer the higher level questions and so finished early and created distractions for the others. When questioned these students reported that the first questions were easy, but the others were "too hard". They either did not understand the vocabulary used in the question but could tackle it after additional teacher assistance, or genuinely did not understand how to work out the solution to the problem. Students felt under some compulsion to "get it all right" and were unable to accept that there might be more than one correct solution or way of tackling the problem, or that their own ideas might be worthwhile, some also felt threatened by the R/EA questions which required them to make judgments not solely based on the musical evidence given. There was a general reluctance to make suppositions even amongst the more able students, which led to feelings of insecurity, unwillingness to attempt the worksheet and a general downturn in motivation. This lowering of motivation caused by attempts to answer higher level questions was an effect which had been noticed previously by Collis, Romberg and Jurdak (1986).

An examination of worksheets set in other subject areas revealed very few questions requiring students to make judgments, so it could be that these feelings could also be due to unfamiliarity with the situation in addition to the inability to tackle these types of questions, and that after further experience in making suppositions and trying to prove them student reactions might be different. To test this theory the same superitem worksheets were then trialled with other students in small group situations. The whole class listened to the music together, then groups of 4-6 students moved apart and discussed the answers to the worksheets. Each student then wrote out their answer in their own words, no copying being permitted within the group, and no collusion with other groups. In the carefully structured heterogeneous groups, the less able clearly saw that no one had a quick solution to the R/EA or even the R questions, and they were all encouraged to think up fresh solutions and explanations. Through this collaboration they were exposed to different styles of thought and their self expectations rose as they realised they understood how to tackle the harder questions. This use of Vygotsky's Zone of Proximal Development (1978) by mixing students at different levels of thinking was a far better classroom strategy for students unused to openended questions, which led to an ostensible increase in motivation and task completion. It was also still possible to identify those students not operating at the Relational level as their lack of comprehension of the intentions of the Relational and Extended Abstract questions could still be seen in their written answers, which were Multistructural lists of factors rather than logical Relational summaries. As due to group collaboration ideas had been pooled so that all students had answered the U and M level questions correctly, it could not be considered a satisfactory technique for the assessment of an individual's performance.

It was also noticed at this stage that the type of music presented influenced the attention paid to it and affected the

number and quality of questions completed. Vocal music caused many students to focuss their attention exclusively on the meaning of the words and they failed to mention the purely musical aspects required to answer the questions.

Stems that contained outline scores or manuscript themes in addition to the listening experience produced more responses at the relational level, probably because it was easier to identify particular passages within the music, and the formal structure could be perceived better with the visual memory aid.

7 : Concert Reviews

All students attended the same concert and heard the same music performed in the same venue by the same performers. A comparison of their reviews of the concert could be expected to reveal the differences in their listening abilities. The students thought that they would have a problem with the length of the concert and find it impossible to remember everything without resorting to extensive notetaking which would detract from the listening experience. It was decided to focus the review on 2 or 3 pieces, which they would describe in detail. This would have been satisfactory if the concert programme had not been changed at the last moment, so that the students could not all focus on the same pieces. Despite this an analysis of their reviews was able to be carried out, which clearly revealed their SOLO levels. Discussion between two students at the concert resulted in very similar responses in terms of the aspects of the music mentioned, but the way they handled the material was different, one stating a succession of facts (Multistructural) and the other reporting on his attitude towards the concert and how the music had created it (Relational). There was obviously scope for collusion between students as they could have collaborated whilst writing their responses as well as whilst collecting the material, so this was not considered a suitable procedure for assessment.

8 : Live Performance Criticisms

These were carried out as part of the normal assessment of practical work in class. Each student performed to the class, who then provided positive feedback by commenting on aspects of their performance which had improved since the last assessment or which needed further attention. Every student including the performer then wrote a short paragraph about the performance.

This procedure was valuable in drawing the attention of the entire class to performing skills that could be refined, and it also provided each student with a record of the occasion that they could refer to in improving their own performance skills. Although it could not be used as a record of their audience listening skills as the observations were shared, the written commentaries did reveal the students' attitudes towards their own practice habits, and several students did make relative judgments and comparisons based on their own observations.

9 : One Word Response Tests

These are commonly used by many Music teachers for assessment of listening, as they can be used with a whole class and the formal test situation ensures individual reponses. They can usefully be used to determine whether the student can identify "What?" they have heard and "What?" the correct musical term for it is, a Unistructural response. This form of testing does not allow for answers to the Relational questions such as "How?" and "Why?", which reveal more about the students' understanding of the music, and is not therefore useful for assessment of the quality of student listening.

10 : Multiple Choice Tests

Multiple choice tests can be constructed which give choices at varying levels of complexity. Unfortunately the preferred answers for Multistructural and Relational questions are usually the longest and therefore easy to spot and select as a guess. The students in this situation are also not required to engage in problem-solving thinking, merely to select between given alternatives. Students operating at the Extended Abstract level are not given any scope and may indeed become discouraged because the structure of the test insinuates that someone else's answer is regarded as more correct than a fresh idea.

11 : Free Response and Essay-type Answers and Tests

Although these type of procedures undoubtedly favour the more literate, they are commonly used as examining techniques, perhaps because they are so versatile in terms of type of question and administration. Questions can be set at any level, to test recall, comprehension and comparison skills or require knowledge to be restructured to fit new situations. They can be administered under formal test conditions, in a classroom, with or without a time limit, or set as homework or assignments. In the formal test situation they can be administered to whole classes, so that there can be no collusion between students, the results can be directly comparable to each other and in the case of music listening it is possible to regulate precisely the conditions under which the music is heard.

Appendix D

Music used in the Test

No 1 Elgar : Bavarian Dance No 3 Op 27

: uses the large twentieth century romantic orchestra
: many changes in instrumentation
: brass led climax in middle section
: tunes that repeat
: simple ABA form in this extract
: standard Western tonality
: driving rhythm
: regular arch shaped melodies
: effective use of full dynamic range available from orchestra
: absolute music but could be programmatic

No 2 Varèse : Ecuatorial (1934)
(extract from a section without the singer)

: small group of orchestral instruments
: instruments used for their timbre on individual notes
: atonal idiom
: many dissonances
: possibly programmatic
: short groups of contrasting rhythmic pitches rather than singable tunes
: long held notes used as both background and foreground
: silences and isolated sounds
: dynamics very precise

No 3 Telemann : Concerto in F for Recorder and Bassoon

: Baroque small string orchestra with harpsichord
: 2 soloists alternating
: orchestra used as accompanist
: contrasts between sound of solists and solo and orchestra
: much melodic, rhythmic and harmonic repetition
: entire section is repeated
: continuous running rhythm and melody
: counterpoint between the two soloists, soloist and orchestra, and within the orchestral parts
: abstract (absolute music), no programme
: dynamics fluctuate only with the number of players

Appendix I

Details of the Test Procedure

Test Procedure : Mini-Lesson

Grade 7 and 8

"Today you are going to hear and write about three pieces of music. Before we start let's go over the different aspects of the music which you could talk about in your answer..." A quick class brainstorm session followed with ideas written on board. When list was complete (teacher added missing aspects) a brief comprehension check of each term was carried out.

List :

Instruments	Repeats of Tunes	Types of Tunes
Contrasts	Harmony/Chords	Speed
Volume	Rhythm and Accents	Mood
Accompaniments		

Ways piece is Like other pieces of music you know
Ways piece is Different from other pieces of music

The grade 7 and 8 students were deliberately given more reminders and help than the older students who had twice as much class time, more background knowledge and were more experienced in listening and discussing music. It was a deliberate attempt to redress the imbalance and create a less threatening atmosphere.

Grades 9 and 10

The list was up on the board as they arrived. Understanding of the meaning of each item on the list was checked and examples from pieces previously used in class were used as illustrations.

List

Orchestration	Timbre	Texture
Formal Structures	Rhythmic Impetus	Tempo
Harmony and Tonality	Articulations	Dynamics
Historical Perspective	Melodic Structures	Imagery
Genre/Style		

Test Instructions (for all grades)

Initial instructions

"First fill in the top of the sheet. This is the first piece so put 1 in the box. Put your pen down when you are ready."

"I am going to play you three different pieces of music and ask you to write about what you hear. You will hear each piece three times with gaps in between. You should have plenty of time to write but if you need extra time you can have as long as you like at the

end. You have been spread out so that you won't be distracted by other people or their ideas. Your ideas are just as important as anyone else's ideas, so write them down confidently. The test will show me what stage you are at in listening so that I can plan our next lessons to help you improve."

" You will hear each piece three times. When you hear it the first time listen carefully, do not write whilst the music is playing, listen and think about the question. I want you to listen very carefully to the music. If you have any ideas you think you might forget, you can of course pick up your pen and jot it down on the sheet, but I'd rather you concentrated on listening and did't write until you'd heard it all the way through. As soon as the music stops jot down your ideas as single words somewhere on the sheet. Think about your answer to the question and start to write as soon as you are ready. After a minute I will play the music again, you may stop and listen or continue writing. After another gap it will be played for the third and final time. I will tell you whether I am playing it for the first, second or third time. After the third time you will have more writing time. If you need extra time you can have it at the end after the third piece or come back later. You can hear the music again later if you want". (No one asked to hear the music again before completing their response, though Piece 3 was requested later.)

"Look at the question" ..read it out.. "remember to prove whatever you say about the music with evidence from what you hear".

"When you have finished put your pen down and I'll know when to go onto the next piece."

The instructions were to be repeated before each piece if necessary.

There were two test papers upon which students were to write their response sheets. The first paper used the instruction : *Describe this piece of music as clearly as possible,* and was used for the Elgar and Telemann pieces (Nos 1 and 3); and the second, used for the pieces by Varèse (piece No 2), was printed with the instruction :
The composer of this music is trying to create a scene or mood in your mind.
In one sentence explain the scene or mood the music creates for you, then explain why the music creates this impression for you.

Test Papers

Listening To Music Year _____

Name_____ Age _____ Grade _____ Month _____

Piece of Music Number [ ]

Describe this piece of music as clearly as possible.

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

Listening To Music Year _____

Name_____Age _____Grade _____ Month _____

Piece of Music Number |__|

The composer of this music is trying to create a scene or mood in your mind.
In one sentence explain the scene or mood the music creates for you, then explain why the music creates this impression for you.

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

Appendix F SOLO Assessment Sheet / Elements Checklist

SOLO ASSESSMENT SHEET

Piece Grade Month Year

Subject Number															
Uni/Multistructural Elements															
Instrument/Voice I.D.															
Group															
Speed															
Volume															
Melodic elements															
Rhythm and Beat															
Mood															
Imagery															
Relational Elements															
Timbre															
Orchestration															
Formal elements															
Structure															
Texture															
Tonality															
Style															
Extended Abstract															
Performance Practice															
Composition Principles															
SOLO Categories															
Prestructural															
Unistructural															
Multistructural a)															
b)															
Relational a)															
b)															
c)															
Extended Abstract															
SCORE															
Percentage															

Appendix G

Responses from the Pilot Study

All responses are presented with their original spellings.
P=Pilot Study E=Elgar V=Varèse T=Telemann

Piece No 1 Elgar : Bavarian Dance No 3 Op 27

Question

"Describe this piece of music in as much detail as possible. You may compare it to other music you have heard or played. You may comment on these aspects of the music : instruments, patterns of tunes, types of tunes, mood, speed and rhythms, harmony and historical background." PE1

Multistructural Responses :

"The piece of music starts at a reasonable pace and then deepens down and climbes higher and higher. The piece of music to me gives the impression of some ballet dancers, dancing under water trying to meet their goal." PE2

Speed, pitch and imagery are the only elements mentioned here.

"The instruments that played mostly throughout the tune was violins and wind, string orchestra and bass drums and a few percussion instruments. The mood was happy and in the middle of the piece, it played into a clamax as if an important event was taking place. The speed and rythm were medium to fast, the rythms were fairly jumpy." PE3

This response adds instrumentation, mood, form and rhythmic aspects, but is still a list and so Multistructural.

Relational Responses :

"The start of this piece is very happy. The tune is jumpy as if imitating a butterfly. The strings play very softly and lightly. The flute seems to imitate the flutter of a butterfly's wings. I can easily picture this butterfly dancing around a garden full of daffodils. The music uses a lot of imitation at the start, playing the same tune with a higher or lower pitch.

The first major change in this piece is the quickening of tempo along with a dynamics change. This could be the butterfly moving faster, or being chased by a bee. Maybe something else dramatic is happening in the garden, that the butterfly is just observing.

To signal the final section, the music is restored to its original mood, using the same instruments. The music starts to flow, not being as jumpy as before - just as if the butterfly is floating down to the ground.

This piece sounds as if it was written in the 19th century, and sounds very "old English". The large use of woodwind gives a very

light feeling to this piece." PE 4
This response combines the pictorial image of the butterfly with the formal structure of the extract to produce a clear picture of the music.

"An orchestra is playing this piece of music. The tune is interchanged between the different sections. The mood of the whole piece is neither depressing or exceptionally lively, but rather a very moderate piece. It begins with the violins playing along and ending each little section of tune with a tiny run. Then they are joined by other instruments. The deeper string instruments such as the double-basses and cellos, play a steady underlying beat at a walking pace. The brass section gradually becomes more noticeable and the volume increases and suddenly dies away. This happens several times until the volume increases to such that the brass overcome the strings and woodwind. During and after this stage the tune and harmony is swaped between the sections and each has small individual parts." PE5
Opening with some general comments, this response then proceeds to describe the extract section by section by combining the main theme of orchestration with other elements.

Piece No 2 Varèse : Equatorial

Question
"In *this piece the composer is trying to create a picture in sound. What do you think his picture is ? Why do you think the music suggests this picture ? Prove you are right by mentioning aspects of the music which led to your decision.*
These aspects may be : instruments and how they are used, patterns of tunes, speed and rhythms, harmony and type of tunes.

As this question required the student to mention an image or mood separately from the discussion of the music, this image was not counted as a musical element when assessing responses to this piece.

Unistructural Responses are characterised by the use of only one aspect of the stimulus. The only musical element in this response is speed :
"This piece is boring. It sounds like a slow death march, with soldiers walking into a battle knowing that they're going to lose. Someone is dreaming about it and he was the only survivor and had to go to all their funerals. At the end it sounds as though he just passes away in his sleep." PV1

Multistructural Reponses to this question were quite short :
"It reminds me of the old horror movies e.g Dracula. The music causes a feeling of suspense not quite sure what is going to happen next. They use piano, drums, trumpet, bass, violins, mostly bass sounding instruments" PV2
Mood and instrumentation are the only elements used to explain the image of the horror movie. Other Multistructural responses were longer.

"It was like Indiana Jones in one of those South American Jungles. The composer creates this picture by mixing high pitched

instruments with low sounding ones. The high ones come in with eerie and high long notes with the bass instruments arriving with loud and rude long notes also. There are also the instruments in the middle, like trumpets and snare drums. Occasionally there are small pauses, then suddenly a loud burst or something. There's really no sort of pattern in this piece or rythym, it just seems to rely on weird sorts of strange notes" *PV3*

This response is trying to explain the unconventional orchestration in this extract but gets confused between the actual notes used and the instruments which play them.

Relational reponses manage to pull separate aspects together, as in this one where orchestration and pitch are seen as contributing to the disconcerting mood :

"It sounds just like music that they would use to try and scare people. It sounds scarey to me as its all disjointed, no instrument ever gets a long go or blends with others. They change from high to low as well, which stops you from adjusting and getting used to the music, and so keeps you scared. At one time the trumpets play very high and then are followed by the percussion who are deep." *PV5*

This Relational response follows a story by direct reference to other uses of similar music and the use of timbre, melodic elements, speed, and volume. It concludes with some speculation as to the effect which would have been created if the music had been orchestrated differently, which could indicate a desire to move into the Extended Abstract level.

"This music creates a picture of a burglary in my mind, because it is similar to music used in spooky movies, where there is someone trying to get into a forbidden area. It makes me see this because at the start it is slow and has a spooky mood created by the bass and muted trumpets. It is just like the start of a burglary, not a lot of drama, the short silences in the music give a high amount of fear inside the felon. As the music quickens up, this gives me the impression that he is closer to his goal. The quick crescendos by the brass seem to be the feeling inside the man as he peeks round each corner and looks behind to cover his tracks.

Without brass this piece of music wouldn't have been possible, The sound of the muted trumpets in particular creates the feeling of uneasiness and tension, and the sudden crescendos wouldn't have had much impact if they had been performed by another instrument." *PV 6*

Piece 3 Telemann : Concerto for Recorder and Bassoon

Question

Describe this piece as clearly as possible.

Mention those aspects of the music which strike you as important from this list : instrumentation, formal pattern of tunes, melodic shape, speed, rhythm amd mood, harmony and historical perspective.

Compare it to similar pieces you have heard, and try to work out who wrote it.

This Multistructural reponse was very short :

"quick, jumpy, loud. fute, obo, recander" PT 1

This is a more comprehensive response using the separate musical elements of mood, rhythm, speed and volume :
This piece is a flowing and happy piece.The speed is quite fast. I think that the rhythm and mood is the most important thing because it sets your mood. Say, the music is a slow and soft piece that seems sad..." *PT 2*

This one is much longer, but is still an unconnected list :
"This piece of music is played by an orchestra, with the violins playing the main part. It is a fast piece in sections and depicts a happy mood. the main tune is made up of many scales, arpeggios and trill. This tune interweaves throughout the string and woodwind sections of the orchestra. the music has the content of the 1700s style of music, and could have been written by Bach or Mozart. The music consists of many phrases each ending very high or low, and slower than the greater portion of the music. It also has many dynamic changes." PT 3

This extract from a Relational response shows the competent integration of the conflict between solists and concertante with the texture of the part writing:
"This piece is played by an orchestra with a duet between a flute and a bassoon, these two have interchanging parts which are tangled together. The orchestra begins the piece but after the first short section plays background when these two begin. Occasionally the orchestra again takes over the lead but is put back in its place by the bassoon who again plays in the duet..." PT 4

Appendix H

Data from the Pilot Study

An analysis of the data revealed that 93% of all responses obtained from the Pilot Study Test were classified at the Multistructural or Relational levels, with half of this number being at each level. The remaining 7% of responses were classified as being at the Unistructural level, and had been stimulated by the second piece, Varèse. As can be seen from Table H.01 all students had at least two responses at the same level, either Multistructural or Relational, and four student response sets were all at the same level. Only 3 students produced a Unistructural level response, and in each case this was for the second piece, Varèse.

Table H.01
Individual Responses Obtained from the Pilot Study Final Test

	Individual Response Patterns				
Main Levels					Total
Multistructural	MMM 1	MUM 3	RMM 2	MRM 2	8
Relational	RRR 3	MRR 2	RMR 1	RRM 1	7

When the responses were sorted into levels and pieces as on Table H.02, it can be seen that there were equal numbers of responses classified at the Multistructural or Relational levels. This table also seems to show differences in the pattern of responses to each of the pieces, and to the second piece, Varèse, in particular. There had also been an expectation that the Varèse would produce the greatest number of Relational responses and although it did produce one more response at this level than either of the other pieces, it was also the source of the only Unistructural responses.

Table H.02
Classification of Pilot Study Responses (actual numbers)

	Pieces				
Levels	Elgar	Varèse	Telemann	Total	Percentage
Prestructural	0	0	0	0	0
Unistructural	0	3	0	3	6.67
Multistructural	8	4	9	21	46.67
Relational	7	8	6	21	46.67

X^2=37.16 at df=4 significant beyond the 0.001 level.

The nonparametric X^2 Test, used for statistical analysis as the SOLO classifications are only on a loose ordinal scale, revealed the significance of the observed difference and individual response patterns were examined to see if the second piece had possibly been responsible for the differences in responses. When the 4 response sets which contained three responses at the same level were discarded, the remaining 11 response sets were classified into groups according to the piece for which the difference in levels occurred, as on Table H.03 below. It could then be seen that for 7

students the level which was different was lower than the others and for 4 students it was higher. Only one student had a different response level for the Telemann, whereas 4 had different levels for the Elgar, and 6 for the Varèse. It seemed as if there might be a difference between responses to the Varèse, but with such a small sample the significance of this was hard to determine. Further work with a much larger sample seemed to be necessary, before any judgments could be made.

Table H.03
11 Difference Level Response Sets Sorted into Groups according to Piece for which the Difference occurred and Type of Difference

	Pieces			
Differences	Elgar	Varèse	Telemann	Total
Higher Difference	RMM 2	MRM 2		4
Lower Difference	MRR 2	MUM 3	RRM 1	7
		RMR 1		
Total	4	6	1	

It can be seen from this table that for 7 students the level which was different was lower than the others and for 4 students it was higher. Only one student had a different response level for the Telemann, whereas 4 had different levels for the Elgar, and 6 for the Varèse.

There seemed to be differences in the pattern of responses to each of the pieces, and to the second piece, Varèse, in particular. This was the only piece with any Unistructural responses and it also had fewer Multistructural level responses and marginally more Relational responses than the others. This might seem to support the idea that the question for this piece would stimulate a wider spread of responses. As the question for this piece required an answer at the Relational level to be correct, there had also been an expectation that this piece would produce the greatest number of Relational responses and it did produce one more response at this level than the other pieces. It seemed as if there might be a difference between responses to the Varèse, but with such a small sample the significance of this was hard to determine. Further work with a much larger sample seemed to be neccessary, before any judgments could be made. Responses to the third piece did not mention its extra length, and students had not become noticably restless during this piece, in fact they seemed to like this piece the most, smiling and tapping the beat as they wrote and asking for it to be played again at their next lesson. It was therefore decided to retain this piece in the test.

Listening Skills

In an attempt to ascertain whether the test was measuring the same skills as the teacher, the responses of six students, who had been identified by their Music teachers as having Listening skills which were well above or below the average were examined and compared to those of the rest of the class. Differences between the three groups are clearly visible on Table H.04, and were confirmed as being significant beyond the .001 level of probability by use of the X^2 Test.

Table H.04
Pilot Study Responses In Groups According to Listening Ability

	Groups		
Levels	Low	Average	High
Unistructural	1	2	0
Multistructural	7	13	1
Relational	1	12	8

X^2=97.25 at df=4 significant beyond .001 level.

As can be seen above, although 8 of the 9 High group responses were at the Relational level, so that this could perhaps be thought to be a characteristic of the students in this group, this level of response was also produced once by a student identified in the Low group for this skill. There were also 12 Relational level responses produced by students who had not been identified as outstanding for their Listening skills, so although those identified for their good listening skills had achieved high levels, it ws not impossible for other students to achieve high levels as well.

Written Fluency

As the test involved a written response, it was expected that those with better writing and linguistic skills would have an advantage over less fluent writers. Students who were exceptionally good or bad in these areas had therefore been identified by their English subject teachers. In this class 4 students were identified as being outstandingly above average, and 3 as being well below average in this area. Once again there was a definite observed difference between the three groups, though due to the high proportion of empty cells this could not be tested statistically.

Table H.05
Pilot Study Responses in Groups According to Written Fluency

	Groups		
Levels	Low	Average	High
Unistructural	3	0	0
Multistructural	6	13	2
Relational	0	11	10

With such a small number of student responses (45) it is not feasible to draw any firm conclusions, but reasons for these differences between the groups could be conjectured. It could possibly be that the Low group of students did not achieve any Relational responses because they were either unable to express their thought coherently in writing and that the other students all had this ability. Written fluency problems could however themselves be related to other language related problem such as comprehension, which would also account for the test response levels, however, though this might be a possible explanation with another group, it was highly unlikely with this particular class who had been trialling this type of test throughout the year. During discussions with the music teachers who taught these students for other areas of the Music course, another possible causal factor emerged, this was motivation towards Music lessons. Each of the 3 Low Written Fluency group students were unwilling to complete tasks set in class, practise their instrument or take part in performances, whereas the High group were conscientious students and enthusiastic band

members, who were continually practising and performing. It seemed as if these two factors, written fluency and motivation had coincided in these students, and it was resolved in the Main Study to identify students separately for written fluency and motivation towards Music lessons in the Main Study and see if they were possibly separate influencing factors.

Performance Ability

Another factor which had been expected to influence the quality of responses was Performance Ability. This was because the ability to listen carefully is essential for the production of a good quality performance. The results of 8 students who had been identified by their instrumental teachers as having either very good or very poor performance skills were therefore isolated and examined. (Although the results of the High group were the same as those of the High groups for Listening Ability, these groups did not consist of exactly the same students.)

Table H.06
Pilot Study Responses in Groups According to Performance Ability

	Groups		
Levels	Low	Average	High
Unistructural	2	1	0
Multistructural	11	9	1
Relational	2	11	8

X^2=113.5 at df=4 significant at the 0.001 level.

Yet again the levels obtained by each group were clearly different, an observation confirmed by the X^2 Test, showing that there could possibly be a relationship between the two skill areas. The high group had mostly Relational responses, and whilst the Low group contained most of the Unistructural reponses found in this part of the study, and every student in this group had at least two Multistructural responses, it was however interesting to note that two of the Low group responses were assessed at the Relational level. So whilst there could be a relationship between the two areas it was also possible that the two skills were not inseparable, and that students who were themselves poor performers could hear clearly and express their thoughts coherently.

School Achievement

As the Pilot Study group had contained 2 students who were outstanding school achievers and 2 who were outstandingly poor in all areas of school life (academic, social, sporting and musical), it was possible to isolate and compare the results of these two groups of students. The observed differences between the results of these two groups of students were striking, although due to the high proportion of empty cells this could not be tested statistically. This observed difference replicated the finding by Biggs and Collis (1982) that SOLO levels showed a close relationship with school achievement.

Table H.07
Pilot Study Responses Grouped According to School Achievement Groups

	Low	High
Individual Response Sets	MUM	RRR
	MUM	MRR

A preliminary investigation of the numbers of students in each grade who were outstanding achievers or failures in all areas of school life revealed that there were usually only 1 or 2 students in each category, and that only some of these students chose to study Music in Grades 9 and 10. So although it would have been possible to identify students for their overall school achievement in the Main Study, it was felt that the numbers would be too small to show any trends that could not be noted in the groups identified for other influencing factors. This characteristic was only noted here because this group of Grade 10s had contained these four exceptional students. The presence of these students also meant that this was not a typical Grade 10 Music class, so that the results obtained from this group might differ from normal patterns, and this was another factor in the decision to conduct another study with a much larger sample of students.

Record Keeping and Self-Assessment Procedures in Classroom

Practical Work Progress Charts
Wall charts with squares for students to colour in when they have proved to another student or adult that they can do the task. As teacher used the chart to identify those needing extra help in class, the students were very honest about keeping this chart accurate, and also used it to find a student they could ask for help before asking for teacher assistance.

Expected standards for School Board Grade 9 and 10 courses : wall charts with list of required tasks for each level and expected standards were used by students to rate their work-in-progress and assess each other before the major assessment periods.

Work-in-Progress Listening

This is like proof reading. It involves listening to a performance or composition before its final version to criticize and offer constructive suggestions for improvement. It can be carried out by students or teachers, and is a normal component of all groupwork in Music.

Appendix I

Music Listening Experiences of Students in the Main Study

Typical Grade 7 and 8 Listening Lesson

Stage One : First hearing and round the class brainstorming.

"What did you hear ?" Emphasis on making a different contribution. No value judgements on contributions at this stage, all ideas apparently accepted with exception of "nice, boring and I like/don't like it", unless reasons for these feelings are given immediately.

Stage Two : Second hearing and preliminary sorting of ideas.

"Listen again and see if you can hear some of the things other people heard the first time". Adjective sheets may be used to stimulate further ideas. Blackboard collation of ideas with some discussion as to suitability, class deciding whether to retain or delete ideas from the original pool...were they true ? Some ideas grouped as teacher writes them up on the board. Additional ideas discussed and added, often teacher adds ideas about form, texture and orchestration and the music may be replayed to show these.

Keywords at this stage : what? when? where?

Stage Three : Facts and Feelings

Relating what was heard and what was experienced on a personal level. Ideas are sorted into 2 lists.
Discussion as to how the facts created the feelings. e.g. "Simon, you said it was a happy piece. Can you pick out some of the facts about the music that made you feel it was happy?"

Keywords at this stage are : how? why? where? in what order?

Stage Four : Recording what we heard.

Copy title and composer from board.
Written work in own words, complete sentences and clear handwriting.
1st sentence : the 3 or 4 facts that seem most important about this piece.
2nd sentence : the feelings the music gives you.
3rd sentence : how the facts created this feeling.
4th ... information about the piece given by the teacher expressed in students own words. (Background, form, reason piece written etc)

Appendix J

Responses from the Main Study

A representative sample of the type of response produced from each piece in the Main Study is given below. All responses quoted include the original spelling, and as far as possible, the original layout. E=Elgar V=Varèse T=Telemann.

Prestructural Responses

Responses which did not attempt to answer the question were classified as Prestructural. As the question for pieces 1 and 3 was *"Describe the music as clearly as possible"* responses which did not mention the music at all were classified at this level. These responses all described scenes which had presumably been stimulated by the music, but as there was no indication of this so they were classified as Prestructural rather than Unistructural.

"someone was running and jumping around the meadow or somewhere,than a bug bad Man or something came chasing it. Then it lost the bad person and went along Dancing and juming around as happy as possible." (E1, Grade 7)

"It is spring. My sister and I are in the garden as big as can be and I keep holdin on her then we see some squrel doing the same and the we see some birds doing the same then we all do it again." (T1, Grade 8)

The next response also fails to mention any aspect of the music, but the events within the story mirror the short episodes within the music. This would be a satisfactory response to a creative writing task based on this music, but it is not adequate as a musical response.

"Two children - brother and sister run into a meadow to play. there is a hill the littel girl stops to pick a daisy and the little boy cartwheels down the hill, the girl runs behind.

At the bottom they play hide-and-seek in the long grass, then make daisy chains, the boy tires of this and throws grass at his sister she begins to chase her brother as they fall down laughing and play leap-frog until the little girl falls over and hurts her knee, and her brother comforts her and wipes away her shortlived tears they jump up to play TAG then race each other back up the hill to where their mother is waiting for them as they happily retreat indoors for lunch." (T2, Grade 9)

The second piece had a two part task where the second part was the one being evaluated. Responses that dealt only with the scene and did not "*explain why the music creates this impression*" were therefore not adequate and were classified as Prestructural.

"I am in a run down old hotel and I am scared I look around the hotel and find a secret passage I entre and fall when I awaken I am sarounded by dead bodys and I am in a pool of Blood I stand and look around I feel like I will be like the other bodys when something grabs my leg I fall to see the bodys comeing closer I fear I shall

die. Once again the dust setals in the hotel." V1, Grade 8)

"I had a picturer of a man running threw a street being followed by police or some monster and he ran into a church for cover.

Or it could be an army man in the batel field walking alone and seeing dead boides lying allover the ground after a bomb hit, and he found his friend deas and turned him over on his back to look at his face." (V2, Grade 8)

Unistructural Responses

Responses which answered the question by using one aspect of the music were classified as Unistructural. Most of these responses mentioned speed, volume or the instruments which were playing.

"It sounds like something is skipping across an oval. People are dancing and they are spinning around getting faster and faster then slowing down. A girl is running and hiding behind trees and her boyfriend is chasing after her. Then he catches her and gives her a big hug." (E2, Grade 8 - speed)

"it is like a muder is gowing to hapen Because the music was loude and like footstepes gowing towords him and men he shot him and ran when Hewas gowing to words his the musick went loude and ckowing and the bloke loucken around an the music quite and thene riely loude when he got the gun and then he sturted to run and some wone chast him and the music went funy and then the music went Bang." (V3, Grade 7 - volume)

"There is some one sneaking around where they are not supposed to be and nearly gets caught. After a while they decide to go somewhere else. The paino sounds like someones foot steps. the cornet sounds like another person looking for him the drums sound like a group of people on patroll." (V4, Grade 8 - instruments)

"This piece of music reminds me of two birds fighting over a piece of bread both trying to grab it and munch into it. Then suddenly a huge crow comes and takes it from them both and the two birds fly away in disgust. Why I said two birds were fighting over it was because every little bit was repeated like a echo." (T3, Grade 10 - melody)

Multistructural Responses

73% of all responses were classified as Multistructural because they used several musical elements to answer the questions. There were differences in the quality of responses at this level, some being very simple whereas others showed a good grasp of musical elements, based their response logically around the structure of the piece and if they had not resorted to using lists would have been classified as Relational.

These responses do not show as much care as some of the Unistructural responses, but they show awareness of more aspects of the music and so are Multistructural :

"piece of music of mood. Volume soft to loud. A bit of speed. Getting faster." (E3 Grade 8)

"voilin, different instruments soft then loud etc. it was'ent like the Boogie stuff. It had different speeds" (E4, Grade 9)

"happy piece orchestra harmony" (E5, Grade 9)

"It sounds like a murder scene it is loud and the music is slow and frightening (it's like the twilight zone)" (V5, Grade 8)

There were also slightly better responses that were presented with more care :

"This peice of music is fast and slow at different time. some of the notes are fast and sharp. It sounds like a waltz. It has no singing. It gets louder and louder as it go's. I has a nice tune. You can hear the violins." (E6, Grade 7 - speed, accent, style, volume, instrumentation)

"This piece is spikey music that you would get in a late night movie. the reason I think this is spookey because It has different Instruments playing little short parts like 2 bars. Some of these instruments are the Piano, Trumpet, organ, and picclo Trumpet." (V6,Grade 10 -instrumentation, melody)

"very quick going up and down the scales very fast with the viloin but not so fast with the tuba always changeing very exciting but it makes you feel relaxed more instruments come in along the way it repeats parts over and over again it is very jerky" (T4, Grade 7 - melody, instrumentation, mood, form, rhythm)

The most common type of Multistructural responses was longer and more detailed like these :

"It is harmony of bass instruments and woodwind. The dynamics change constantly, there is a big build up and then it goes softer. It is rather fast, has a jumpy melody, quite fast. There are lots of different instruments playing at the same time. Starts off at a reasonable pace gets faster and louder. It's like both Brass and Woodwind are competing against each other. Some parts it is tongued others it is slurred." (E7, Grade 9 -instrumentation, volume, melody, speed, articulation)

"it sounds like a murdering peice of music and like someone has comited a crime. The music is a very deep peice of music. It had alot of bass sounds. It creates this impression because of the low sounds at the start of the music and then starts to become pretty fast. and when the snare drums come in it sounds like apiece of war music. And when the whole lot of the instruments get to gether they give us a impression of a hole lot of different things like war, action scary murder. I recon its a good piece of music for someone who is doing a play with all these things in it. It is a drama like peice of music

I thought it was alright" (V7, Grade 7 - pitch, speed, instrumentation)

"My thoughts on this piece of music is that it was very fast and it was kept at the one leval of speed to me. I think that it is a very hard piece of music and it most bumpy and loppy. It sound Very lively and whispy. It is a very high piece of music. This piece of music had a flute in it. I think it would be played at a Ballet or some kind of fast dance." (T5, Grade 8 -speed, melody, mood, pitch, instrumentation, style)

"Very quick tempo. The scales go up and down all the time. Each instrument has their go at separate times. There didn't really seem to be a change in dynamics. this piece of music didn't seem to build up to anything. that song gave the feeling of happiness and joy. That was the only one out of the three that i didn't like. Everything in that piece of music was the same. I found it boring. Actually it also sounded like 2 old men rambling on about somthing that wasn't of great interest to anyone." (T6, Grade 10 - speed, melody, orchestration, volume, structure(?), mood, imagery)

There were also some competent Multistructural responses that were mostly logically based around the formal structure of the pieces and seemed to be almost ready to make the transition to the relational level. They still seemed to be making lists, so they were classified as multistructural :

"This piece reminds me of a war. It starts off at a medium pace, comes to a FF climax in the middle and finishes piano and gently like someone has died.
The music stays at a fairly constant tempo. At the climax timpanis are used as well as bass instruments. Flutes and strings are used to create a feeling of falling.

There are two main lines , the melodic line played by the violins and some woodwind instruments, and the bass line which is the beat of the piece mainly." (E8, Grade 10)

"At the beginning the instruments or instrument went on for long periods of time with the same note played continuously while some other instruments went on playing other tunes before the other one stopped. It was noisy and weird The volume level kept changing and sometimes it had a real shrill instrument playing and abruptly went to a very low instrument The speed kept changing. At the beginning it had short 'blasts' of music from a piano and other instruments and it gradually got louder. Really, you could say it was mostly slow, but mostly loud. It probably felt eery and jumpy, as if it wanted to make you calm then make you scream.
It wasn't much like other pieces of music I've heard, because other music mostly has music flowing from one bar to another. This one was 'jotty' as it kept 'leaping'." (V8, Grade 7)

"This piece of music is <u>VERY</u> fast, it's lively, and it jumps about alot, the violins have the tune mostly, but after a while a horn or what sounds like a horn but I'm not sure if it is a horn, takes the tunes and going on a bit, and after that a recorder, or what sounds like one, has the tune, it has a harmony line, and different instruments play that, so it varies it tunes and harmony lines. There is also repeats of the beginning, and of different parts in this particular piece of music. I think the composer must have had a hard time writing this music onto paper it must've taken hours too do. It has lots of trills in it, not exactly hundreds but quite a lot. Overall, this piece was pretty good to listen too." (T7, Grade 7)

<u>Relational Responses</u>

Although all Relational responses created logical arguments by basing their description on one relational element such as form or style and using other elements as illustrations or evidence for their thought, these responses also varied in their quality. Each of the responses at this level was based on one aspect of the music, usually the formal structure of the piece, and the difference in quality lay in the way other elements were integrated. This Grade 7 response focusses on the changes in volume which characterised the sections within the Elgar extract :

"This piece of music had a very unusual style. It would go loud one time and soft the next time, sometimes it went extremely soft. The violins played a real fast noticable tune and the louder parts were played by the larger instruments. It was jumpy music. When it was in the soft parts it made me feel happy, But in the loud bits it had a scary feeling like someone being chased. I think it was played by the whole orchestra because it varied so much in volume and

tune." (E9, Grade 7)

The structure of the imagined story determines the form of these responses to the Varese extract :

"A small cartoon boy is walking slowly through a very dark cave, he turns a corner and sees a very large diamond, he touches it, it makes him imagine he is an army person and is walking through a hot, moist jungle.

The music makes me feel like this because at first it is soft and spooky. then it begins to get some excitement in it, this is when the boy sees the diamond. Then, as he approaches the music goes slower. As he touches it the music "Booms". Drums sound and it makes the imagination of an army come into my head. the type of music that is played in movies when someone is going through a jungle is played. It is like a jungle scene." (V9, Grade 7)

"The beginning of the piece is hesitant and detached. It forms an image in my mind of a criminal lurking around a deserted alleyway, constantly watching out for the police.

The detached notes at the beginning and then the long high pitched notes indicate the criminal tramping around the alleyway, but still being observant, looking out for the police.

The sudden crash of the bass drums and then the increase in speed and volume indicate the criminal has been spotted, and he begins, desperately, trying to run along the cold deserted alleyways to keep in front of the police. As the piece reaches a descrescendo at the end, the effect of the gradual softness indicates the criminal has managed to disappear into the distance, never to be seen again..... " (V10, Grade 10)

This response to the Telemann Concerto is centred around the contrapuntal use of melodies :

"This piece of music is very fast and lively. It starts off with the violins playing a figure that is very light and jumpy. Then an oboe plays a figure like what the violins played that is just about as jumpy. the tune is then played by the lower strings. Underneath the tune is a bass line that created the jumpiness because it is very staccato. this is played by the strings. After this the tune switches to a piccolo that make it feel very 'light and fluffy' because it is such a high instrument.

The whole way through, the piece is fairly fast and their aren't any dramatic dynamic clashes." (T8, Grade 10)

More sophisticated Relational responses integrated many musical elements effectively to form a coherent argument, and whereas these were usually found in Grade 10 responses, they did occur in Grade 7 responses :

"This piece of classical music started off with tune A. Tune A was slow pieces with the strings. While they kept the tune, the violins would flutter around with quavers. The main tune would lead up to something then the violins would play their part. It reminded me of autumn leaves fluttering to the ground sometimes fast sometimes slow, depending on the wind, throughout the tune. Tune B was leading up to something, getting louder, Then tune C hit me. It was very fast, with booms from the drums adding to the excitement. Tune A came back again and the coda was simply the main tune slowed down and quieter. Even though it was slower, the composition finished with a crotchet." (E10, Grade 7)

"This piece of music starts with violins playing a fast, soft tune. In the background a deeper instrument like a cello or a doubld bass can be heard thumping away. This tune is a very lively tune

which jerks around at one time it becomes loud and thumping but this only lasts a few seconds and it is soon back to its usual fast pace. This pace lasts for a while and reminds me of a volvo or another expensive car ad. It gets faster but like before it didn't last long and soon it was back to the fast soft tune. This tune stops and a repeatitive short tune comes in. This glides into the old tune which is played nearly all the way through. This gets loud then soft making its way to its finale." (T9, Grade 7)

"The piece is played by an orchestra in a moderate 2, though it moves along fairly quickly. The piece starts with the tune alternating between the wind and strings at moderate volumes and creates a very light, jovial atmosphere, almost like a novelty piece. The tune itself flows along in an almost continual stream of notes and is followed by a large loud build up into the next section, which takes on a darker more serious tone. In parts this section goes extremely fast and after the initial stream of descending passages, a quieter, more peaceful tune emerges played mainly by the wind which finishes the music quietly."(E11,Grade 10)

"This piece is played by a full orchestra. The speed is moderate at the beginning, there is a light flowing tune carried by the strings (violins) and woodwind (possibly oboe and/or clarinet) in turn. The brass and flute have fill ins. the accompaniment is rich sounding and strong played by brass and possibly double bass. The accompaniment brings in a crescendo and seems ot carry the rest of the orchestra to a peak and a second tune. The crescendo brings the momentum up to a point, and seems to drop off, strting a tune similar to that at the beginning. The violin, glockenspiel and flute play the tune in turn. The dynamics vary greatly in volume and articulation. The piece is in the form ABA" (E12, Grade 10)

"The music creates a dark, morbid atmosphere which creates a feeling of fear and suspense; perhaps ultimately death. The two main contributing factors to this atmosphere are probably the selection of instruments and the abstract, unpredictable tune which doesn't give the slightest clue to what's going to happen next. The use of the piano and trumpet at the start with the detached tune and minor feel set the atmosphere which is added to by the occasional, sometimes unexpected percussion. The section where the trumpets crescendo and stop suddenly gives the impression of a climax to some event. the music also tends to create a graveyard type, supernatural effect which is added to by the use of a whistle or a flute to create a owl's hoot. The organ also seems to support this idea as it is very like an organ church which is often used for this very effect." (V11, Grade 10)

Appendix K

Data from the Main Study

The Retest Data for each Grade

An examination of the retest data for each grade can give a more detailed picture of trends. From Grade 7 to 8, the number of responses at the Unistructural and Relational levels remained similar, whilst the numbers at the other levels altered, as can be seen on Table K.1. There was a decrease of 11 in Multistructural responses, and whereas there were 2 more Relational responses in Grade 8, there were also 10 Prestructural responses compared to none in the first test taken in Grade 7. As this was the students' second test the Prestructural responses cannot be attributed to lack of comprehension of the question, and could perhaps be due to another factor such as a decrease in motivation. When categories were combined to eliminate cells which were too small for analysis, no significant difference between the groups was found.

Table K.1 Classifications of Responses by the 48 Students Tested Twice in Grades 7 and 8

	Grade in which Test was Taken	
	First Test	Second Test
Levels	Grade 7	Grade 8
Prestructural	0	10
Unistructural	16	15
Multistructural	127	116
Relational	1	3
Total	144	144

With categories PU and MR combined, $X^2=1.64$ at df=1. Not significant.

Changes in response levels from Grade 8 to Grade 9, recorded on Table K.2, also show differences, although this was a rather small group of only 12 students. Once again there was the appearance of Prestructural responses (2) on the retest which had not been found on the first test. There was also a small increase in the number of Unistructural responses and a larger increase in Relational responses. The difference between the two groups was significant beyond the 0.001 level using the X^2 test, where X^2 was calculated at 24.07.

Table K.2 Classification of Responses by the 12 Students Tested Twice in Grades 8 and 9

	Grade in which Test was Taken	
	First Test	Second Test
Levels	Grade 8	Grade 9
Prestructural	0	2
Unistructural	3	6
Multistructural	29	17
Relational	4	11
Total	36	36

With categories PU combined to remove the empty cell, $X^2=24.07$ at df=2 significant at 0.001 level.

Analysis of the responses from the 14 students tested in both Grades 9 and 10 shown on Table K.3 revealed no lower level responses at all. The null hypothesis that there was no significant difference between the two tests was upheld by the X^2 test, which revealed that the apparent differences had a 1 in 10 probability of occurring by chance.

Table K.3 Classification of Responses by the 14 students Tested Twice in Grades 9 and 10

	Grade in Which Test Was Taken	
	First Test	Second Test
Levels	Grade 9	Grade 10
Prestructural	0	0
Unistructural	0	0
Multistructural	9	16
Relational	33	26
Total	42	42

$X^2=2.8$ at df=1. Not significant.

The 13 Grade 10 students who were tested and retested in Grade 10 were, apart from the Grade 9 Retest Experiment group, the only students to have been retested less than a year after their initial test. They might therefore be expected to show some effects due to test familiarity. However the responses from these students on both tests were so similar, as can be seen on Table K.4, that the X^2 test revealed that the slight differences observed were probably due to chance.

Table K.4 Classification of Responses by the 13 students Tested Twice in Grade 10

	Grade in Which Tests Were Taken	
	First Test	Second Test
Levels	Grade 10	Grade 10
Prestructural	0	0
Unistructural	0	0
Multistructural	9	6
Relational	30	33

$X^2=0.76$ at df=1. Not significant.

Relationship between Influencing Factors and SOLO Levels

As the comparison of high and low groups for each factor had revealed general differences in the SOLO levels achieved by each group, though not necessarily by individuals within each group, it was considered necessary to examine the characteristics of students

who had achieved both high and low SOLO levels. As the majority of responses were at the Multistructural level, and most students who produced Prestructural, Unistructural and Relational level responses also produced responses at this level, it was decided to focus on students who produced responses at these less common levels. To avoid duplication and present as accurate a picture as possible, only responses from first tests were considered.

As can be seen on Table K.5 seventeen students produced at least one Prestructural level response on their first test. Of these students, 11 had not been identified as outstanding for any factor, and are listed on the table as the Average group, but 6 had been identified, 5 in Low groups and 1 in the High group for Performance Ability.

Table K.5 Characteristics of Students with Prestructural Responses on First Tests

	Identification Groupings			
Factors	Low	Average	High	
Music Listening	1	16	0	
Written Fluency	3	14	0	
Motivation	3	14	0	
Performance Ability	3	13	1	
[Total No of Students	5	11	1	17]

11 students with Prestructural responses were not identified for any factor.

Most of these students had only been identifed for one factor, but 1 student had been identified in the Low group for all four factors and 2 had been identified for 2 factors. These are shown on Table K.6.

Table K.6 Multiple Identifications of Students with Prestructural Responses on First Tests

	Identification Groupings	
Factors	Low	High
4 factors	1	0
3 factors	0	0
2 factors	2	0

A comparison was also made with the total number of students identified for each of these factors to assist in assessing the importance of each factor in determining a Prestructural response. This comparison is presented in Table K.7. In each case the percentage of students with Prestructural responses was small in comparison to the numbers of students identified for these factors but producing higher level responses. One student had been identified as having poor Music Listening skills, which accounted for 10% of the group of 10 students identified for poor Music Listening skills. Written Fluency and Motivation each with 3 students in the Low group also accounted for a small percentage of students identifed for poor skills in these areas. The one factor that might be expected to have the least influence on the test, Performance Ability, was the one where the largest percentage of students identified in the Low group for this factor had Prestructural responses, however this 15% was still only made up of 3 students. Evidence of strong influences on Prestructural responses seemed to be lacking here, so this examination was extended to include all first tests with Unistructural responses.

Table K.7 Characteristics of Students with Prestructural Responses on First Tests as a Percentage of the Total Number of Students Identified for Each Factor

	Identification Groupings	
Factors	Low	High
Music Listening	10	0
Written Fluency	13.05	0
Motivation	10.72	0
Performance Ability	15	4.17

As can be seen on Table K.8, 86 students had been assessed as having produced Unistructural responses on their first tests. Of these, 56 were not identified for any factor, whereas 30 had been identified for at least one of the possible influencing factors with 25 students identified in Low groups and 5 students in High groups.

Table K.8 Characteristics of Students with Unistructural Responses on First Tests

	Identification Groupings			
Factors	Low	Average	High	
Music Listening	8	78	0	
Written Fluency	15	70	1	
Motivation	17	66	3	
Performance Ability	13	69	4	
[Total Number of Students	25	56	5	=86

56 students with Unistructural responses were not identified for any factor.

Of the 25 Low group students, 16 had been identified for more than one factor. As can be seen on Table K.9, two students in the High groups had also been identified for more than one factor.

Table K.9 Multiple Identifications of Students with Unistructural Responses on First Tests

	Identification Groupings	
Factors	Low	High
4 factors	3	0
3 factors	6	1
2 factors	7	1

Table K.10 shows that the majority of Low group identifications for all four possible influencing factors had come from students who had produced Unistructural responses. The 8 students with Unistructural responses who had been identified as having exceptionally poor Listening Ability accounted for 80% of those in the Low Listening Ability group. Furthermore as, when added to those attaining the Prestructural level, shown on Table 6.35, this figure rose to 90% of the students identified in the Low group for this factor, it could be postulated that poor Listening Ability was a factor in determining low SOLO levels. However as these percentages were only created by 9 students, whereas 11 students produced Prestructural responses and 56 students produced Unistructural responses but were not identified in this Low group, it cannot be asserted from this evidence that poor Listening Ability was the only, or even the major reason, for these levels of response. The picture with the other three factors is even more confused as responses at this level were produced by both students identified as having both outstandingly weak and outstandingly strong skills in these areas.

Table K.10 Characteristics of Students with Unistructural Responses on First Tests as Percentage of Students Identified for Each Factor

	Identification Groupings	
Factors	Low	High
Music Listening	80	0
Written Fluency	65.22	5.27
Motivation	60.72	9.68
Performance Ability	65	16.67

A consideration of the combined characteristics of students with these low level responses on their first tests, Table K.11, is not much more revealing. The total number of students with Prestructural and Unistructural responses was 103, but as 77 (75%) of these students had 2 responses at the Multistructural level, this figure should not be regarded as an indication of a large number of students performing at the lower levels. The 37 students with Prestructural and Unistructural responses who were identified for influencing factors also accounted for only 35% of these 103 students, so the majority of Prestructural and Unistructural levels responses (65%) still could not be attributed to any of these factors.

Table K.11 Characteristics of Students with Pre- or Unistructural Responses on First Tests as a Percentage of the Students Identified for Each Factor

	Identification Groupings	
Factors	Low	High
Music Listening	90	0
Written Fluency	78.26	5.27
Motivation	71.43	9.68
Performance Ability	80	20.84

The characteristics of students who had produced Relational responses were examined in the expectation that they would be found amongst the High groups for the possible influencing factors. As can be seen on Table K.12, this expectation was not unfounded, but there were still 34 students with Relational level responses who had not been identified for any factor, a greater number than the 20 identified in the High groups. 55 students produced Relational level responses on their first tests, and of these 20 were identified in High groups. It was interesting to note that one student identified by teachers to be in the Low groups for 2 factors, Written Fluency and Motivation, had achieved a Relational level response. For this one student these factors had not proved to be a handicap. There were also 34 students who had produced Relational level responses but who had not been identified for any factor, which is a greater number than the 20 students identified in the High groups. It would therefore seem to be imprudent to attribute the ability to produce a Relational level response to these factors.

Table K.12 Characteristics of Students with Relational Responses on First Tests

	Identification Groupings			
Factors	Low	Average	High	
Music Listening	0	48	7	
Written Fluency	1	43	11	
Motivation	1	43	11	
Performance Ability	0	41	14	
Total Number of Students	1	34	20	55

34 students with Relational responses were not identified for any factor.

Of the 22 students who were identified in either High or Low groups, 13 had been identified for more than one factor. As can be seen on Table K.13 below, the one Low group student had been identified in two Low groups and 12 of the High group students had been identified for two or more factors.

Table K.13 Multiple Identifications of Students with Relational Responses on First Tests

	Identification Groupings	
Factors	Low	High
4 factors	0	4
3 factors	0	3
2 factors	1	5

Once again a comparison was made between those who had achieved Relational level responses and had been identified for these factors and those who had not been identified, and the percentage of each group attributed to the Relational response students is shown on Table K.14. As this table shows, the majority of those identified in the High groups for Written Fluency and Performance Ability had achieved Relational level responses, and a large percentage of those identified for outstanding skills and Music Listening and good Motivation also achieved responses at this level. There were however still many students who had produced Relational level responses but had not been identifed as outstanding on any of these factors.

Table K.14 Characteristics of Students with Relational Responses on First Tests as a Percentage of the Students for Each Factor

	Identification Groupings	
Factors	Low	High
Music Listening	0	43.75
Written Fluency	4.35	57.90
Motivation	3.58	35.49
Performance Ability	0	58.34

Of the 55 students with at least one Relational level response on their first test, 19 had produced fully Relational response sets (RRR) from their first test. None of these students had been identified in Low groups but 9 had been identified as having outstandingly good skills in at least one of the factors thought to be influences on test results. As can be seen on Tables K.15 and 6.40, these 9 students were identified 18 times, but although they were undoubtedly talented they made up a small percentage of the

total number of students isolated for each of the factors. Of these 19 RRR students 10 had not been identified for any possible influencing factor, and 75% of the students identified for outstandingly good skills in musical performance, for instance, were not amongst these consistently high SOLO level performers. It does not therefore seem feasible to attribute high SOLO levels solely to these factors.

Table K.15 Characteristics of Students with 3 Relational Responses on First Tests

Factors	Students	
	Number of students	Percentage of High Group
Music Listening	2	12.5
Written Fluency	4	21.05
Motivation	6	19.36
Performance Ability	6	25
No factors identified	10	0

Appendix L

Retest Experiment

November/December 1988 Grade 9

Content of Listening Lessons between Tests

Lesson 1

"Water Under Snow is Weary" by Finnish composer Harri Wessman was listened to and discussed. This piece is in two sections, the first being a virtouso flute and piano duet and the second being a verse and chorus type song for children's part choir accompanied by flute, piano and string quartet.
As it was sung in Finnish, the words were not a distraction. As the sections are completely different, it was hoped that formal structure would be easily recognised.

The elements of music were sorted into two categories : basic fact and relational ideas.

Basic Facts
Instruments Groups of instruments
Speed Volume
Melodic ideas - tune direction - runs - jumps - pitch range
Rhythm - pulse - time - accents - stylistic characteristics
Mood - why ?
Imagery - picture - story - why ?
Relational Ideas : Timbre/tone colour
Orchestration - groupings and tone colours created by them
Formal Structure - patterns of tunes/sections - genre - type of piece
Texure - homophonic - polyphonic - tune and accompaniment - imitation - composing devices
Tonality - key/scalic system - harmony - concords/discords
Style - type of music - period

It was explained that as all the relational ideas implied knowledge of the basic facts they were higher level concepts and all Level Three students should be using them in a logically connected way. Written work which did not use these ideas or relate them to each other would be assessed at level Two.
Some students decided to make notes about these concepts.

All students participated in this first discussion and were given 10 minutes to write about this piece of music. Their writing was given Levels to show the level at which they were operating and the reasons for these assessments were discussed individually with those who wanted more detailed information. (In general those students already operating at a relational level asked very few questions. Further explanation was requested by those having difficulty with formal and textural elements.)

Music used in subsequent lessons for class debate and writing :

Expletive Delighted : Fairport Convention
Can Can : Offenbach
Five Pieces for Orchestra : Webern
La Rejouissance : Handel

These pieces were selected for their demonstration of formal structures, texture, orchestration, style and period and lessons focussed on these elements in order to increase awareness of the higher level musical concepts. Baroque and modern compositional techniques were discussed through the Handel and Webern, which have similarities to the Varèse and Telemann test pieces. A programmatic or romantic orchestral piece similar in style to the Elgar test piece was deliberately not included as it was felt that it might stimulate debate on imagery rather than on more intrinsically musical elements.

During the time between tests the class were also working on individual compositions and were involved in work-in-progress listening and criticism. The task they had been set was to create, perform and record a 1-3 minute long piece with melody and harmony which had 3 sections with recognisably different moods. As part of the composition process they requested extra help with techniques of melodic manipulation and these ideas were discussed, demonstrated and experimented with :

alteration of time signature
alteration of keys (adding # and bs)
changing the articulation and accents
phrasing
ornamentation
harmonic effect (7ths, minors, added notes, chords of 2nd)
pitch (octave shifts)
volume (terrace dynamics v. use of crescendo)
speed
imitation - sequences - motivic alteration (upside down)

These concepts were always dealt with as factors which could create musical effects and moods. It was felt that this short period of intensive thinking and listening to their own compositions might possibly affect the student's reactions to other music.